CENTRE
HENRI LEBESGUE
CENTRE DE MATHÉMATIQUES

2022
29 JUIN
1ER JUILLET

IA ET SANTÉ

APPROCHES INTER DISCIPLINAIRES

NANTES

EN PARTENARIAT AVEC
SOCIÉTÉ FRANÇAISE DE STATISTIQUE, GDR STATISTIQUES ET SANTÉ, PROGRAMME DATASANTÉ

COMITÉ ORGANISATEUR

LISE BELLANGER
NANTES UNIV.

DAVID CAUSEUR
INSTITUT AGRO - RENNES-ANGERS

MATHIEU EMILY
INSTITUT AGRO - RENNES-ANGERS

VALÉRIE GARÈS
CASA RENNES

PIERRE-ANTOINE GOURRAUD
NANTES UNIV. - CHU NANTES

DIANA MATEUS
NANTES UNIV. - CENTRALE NANTES

BERTRAND MICHEL
NANTES UNIV. - CENTRALE NANTES

FRÉDÉRIC PROÏA
UNIV. D'ANGERS

AYMERIC STAMM
CNRS LMJL

STÉPHANE TIRARD
NANTES UNIV.

COMITÉ SCIENTIFIQUE

LISE BELLANGER
NANTES UNIV.

DAVID CAUSEUR
AGROCAMPUS OUEST

MICKAEL GUEDJ
HANDIBIOTEX

NOLWENN LE MEUR
INSERM RENNES

NICOLAS MOLINARI
UNIV. DE CHU DE MONTPELLIER

EMELINE PERTHAME
INSTITUT PASTEUR PARIS

CÉCILE PROUST-LIMA
INSERM - UNIV. BORDEAUX

NATHALIE VIALANEIX
INRAE TOULOUSE

WWW.LEBESGUE.FR

PARTNERS
INSTITUT DE RECHERCHE MATHÉMATIQUE DE RENNES
LABORATOIRE DE MATHÉMATIQUES JEAN LERAY
DÉPARTEMENT DE MATHÉMATIQUES, ENS RENNES
LABORATOIRE DE MATHÉMATIQUES DE BRETAGNE ATLANTIQUE
LABORATOIRE ANGEVIN DE RECHERCHE EN MATHÉMATIQUES

SUPPORTS
AGENCE NATIONALE DE LA RECHERCHE
RÉGION BRETAGNE
RÉGION PAYS DE LA LOIRE

AFFILIATIONS
UNIV. DE RENNES 1
UNIV. RENNES 2
NANTES UNIVERSITÉ
UNIV. D'ANGERS
UNIV. BRETAGNE SUD
UNIV. DE BRETAGNE OCCIDENTALE
CNRS
INSA RENNES
INRIA
ENS RENNES

Table of contents

Artificial intelligence and health: interdisciplinary approaches

From June 29th to July 1rst 2022 - Nantes

The conference «Artificial Intelligence and health: interdisciplinary approaches» will be held in Nantes at the MSH Ange Guépin from 29 June to 1 July 2022. This scientific meeting is part of the thematic semester 2022 «Machine Learning / Artificial Intelligence» organized by the Henri Lebesgue Mathematics Center in partnership with the Data Santé program and the GdR "Statistics and Health". It is sponsored by SFdS. It will offer a unique opportunity to exchange knowledge, share ideas and best practices around the theme «AI and Health», favoring interdisciplinary approaches. It brings together researchers and practitioners, from public institutes or the private sector, who contribute or use AI approaches in the health field.

Here is a non-exhausting list of yet predominant topics that the conference aims at covering:

- Statistical learning
- E-Health
- Data science and health
- Digital epidemiology
- Big data and public health
- Statistics and personalized medicine
- Health economics
- Health and human and social science
- Modeling of infectious diseases
- Large-scale clinical research - Integration of omics data
- Large-scale clinical research - Integration of imaging data
- Drug target
- …

Committees

Steering committee

Lise Bellanger (LMJL-Univ Nantes, Nantes)
David Causeur (IRMAR-Agrocampus Ouest, Rennes)
Mathieu Emily (IRMAR-Agrocampus Ouest, Rennes)
Valérie Garès (IRMAR-INSA, Rennes)
Frédéric Proïa (LAREMA-Univ Angers, Angers)
Aymeric Stamm (LMJL-CNRS, Nantes)

Local organizing committee

Lise Bellanger (LMJL)
Pierre-Antoine Gourraud
(clinique des données, ITUN - CRTI - UMR INSERM 1064 – CHU Nantes)
Diana Mateus (LS2N, ECN)
Bertrand Michel (LMJL, ECN)
Stéphane Tirard (resp du projet Data Santé, Centre François Viète)

Scientific committee

Lise Bellanger (LMJL-Univ Nantes, Nantes)
David Causeur (IRMAR-Agrocampus Ouest, Rennes)
Mickael Guedj (Nanobiotix, Paris)
Nolwenn Le Meur (EHESP, Rennes)
Nicolas Molinari (CHU, Montpellier)
Emeline Perthame (Institut Pasteur, Paris)
Cécile Proust-Lima (INSERM, Bordeaux)
Nathalie Vialaneix (DR INRAE UR875 MIA-T Toulouse)

Contacts

Administration: secretariatlmjl@univ-nantes.fr

Lise Bellanger: lise.bellanger@univ-nantes.fr
mobile: 06 75 14 46 07

David Causeur: david.causeur@agrocampus-ouest.frww

Mathieu Emily: mathieu.emily@agrocampus-ouest.fr

Valérie Garès: Valerie.Gares@insa-rennes.fr

Pierre-Antoine Gourraud: pierre-antoine.gourraud@univ-nantes.fr

Mickael Guedj : mickael.guedj@gmail.com

Nolwenn Le Meur: Nolwenn.LeMeur-Rouillard@ehesp.fr

Diana Mateus: diana.mateus@ec-nantes.fr

Bertrand Michel: Bertrand.Michel@ec-nantes.fr

Nicolas Molinari: n-molinari@chu-montpellier.fr

Emeline Perthame: emeline.perthame@pasteur.fr

Frédéric Proïa: frederic.proia@univ-angers.fr

Cécile Proust-Lima: cecile.proust-lima@u-bordeaux.fr

Aymeric Stamm: aymeric.stamm@math.cnrs.fr

Stéphane Tirard: stephane.tirard@univ-nantes.fr

Nathalie Vialaneix: nathalie.vialaneix@inrae.fr

Program

Wednesday June 29th 2022

8h30-9h30	Reception & inscription	
9h30-10h00	Opening of the conference (Amphitheater Simone Weil)	
10h00-11h00	Invited speaker 1 : **Grégory Nuel** (LPSM-CNRS 8001, Sorbonne Université) *Inferring causality from a mixture of observations and interventions* (Amphitheater Simone Weil) Moderator: Mickael Guedj	
11h00-11h30	Coffee break	
	Contributed session 1 (Amphitheater Simone Weil) Moderator: Emeline Perthame	**Contributed session 2** (Room B) Moderator: David Causeur
11h30-11h50	**Antoine Bichat** *Analysis of Cancer Dependency Maps*	**Pierre Drouin** *Classification semi-supervisée de séries temporelles de quaternions pour l'analyse des troubles de la marche dans la sclérose en plaques*
11h50-12h10	**Zeno Loi** *Topological machine learning approach for linkage exploration in transcriptomes*	**Eloïse Inacio** *A coarse-to-fine needle extraction algorithm for the modelisation of electric field in electroporation ablation*
12h10-12h30	**Bastien Chassagnol** *Robust deconvolution of transcriptomic samples using the gene covariance structure*	**Hava Chaptoukaev** *Assessing Multiple Imputation of Missing Values for Robust Analysis of Telehealth Kiosk Data*
12h30-14h30	Lunch break – Poster session (Room A)	
14h30-15h30	Invited speaker 2 : **Basile Chaix** (iPLesp Sorbone Université) *The use of wearable sensors and smartphones in environment - health research* (Amphitheater Simone Weil) Moderator : Nolwenn Le Meur	
15h30-16h00	Coffee break	
	Invited session 1 : Multi-source data integration (Amphitheater Simone Weil) Moderator: Valérie Garès	
16h00-16h30	**Anne-Sophie Jannot** (University of Paris Cité and Georges Pompidou European Hospital) *Reuse of the French National Health Insurance data for patients suffering from rare diseases: the Dromos project challenges*	
16h30-17h00	**Eric Letouzé (**CRCI2NA, Inserm UMR 1307, Nantes Université) *Multi-omics data integration in cancer research*	
17h00-17h30	**Erwan Drezen** (entreprise CUBR) *Better data for a better science: the challenge of linking health databases*	
18h00-19h30	Cocktail hall MSH	

Thursday June 30th 2022

8h30-9h00	Reception & inscription	
9h00-10h00	Invited speaker 3 : **Rodolphe Thiébaut** (INRIA - INSERM – CHU de Bordeaux - Université de Bordeaux - ISPED) *Analyzing transcriptomics data for understanding and predicting vaccine response in clinical trials* (Amphitheater Simone Weil) Moderator: Cécile Proust-Lima	
10h00-10h30	Coffee break	
	Invited session 2 : Neuroscience and image analysis (Amphitheater Simone Weil) Moderator: Emeline Perthame	
10h30-11h00	**Alexandre Gramfort** (Inria, CEA, Université Paris-Saclay) *Machine learning without human supervision on neuroscience data*	
11h00-11h30	**Florence Forbes** (Inria, Universite Grenoble Alpes, equipe Statify, Laboratoire LJK) *Simulation based inference for high dimensional inverse problems: application to magnetic resonance fingerprinting*	
11h30-12h00	**Christophe Zimmer** (Computational Biology Department, Institut Pasteur, Paris) *Deep learning for biomedical imaging*	
12h00-13h30	Lunch break – End of Poster session (Room A)	
13h30-14h00	Presentation of the best poster award	
14h00-15h00	Invited speaker 4 : **Stéphanie Allassonnière** (Université Paris Descartes & Ecole Polytechnique) *Data Augmentation in High Dimensional Low Sample Size Setting Using a Geometry-Based Variational Autoencoder* (Amphitheater Simone Weil) Moderator: Bertrand Michel	
15h00-15h30	Coffee break	
	Conributed session 3 (Amphitheater Simone Weil) Moderator: Nolwenn Le Meur	**Conributed session 4** (Room B) Moderator: Lise Bellanger
15h30-15h50	**Anthony Devaux** *Random survival forests for competing causes with multivariate longitudinal endogenous covariates*	**Cheïma Boudjeniba** *Identification of consensus whole blood transcriptomic gene modules in patients with primary Sjögren' s Syndrome*
15h50-16h10	**Audrey Lavenu** *Comparaisons de méthodes pour données de survie en grande dimension sur de petits échantillons*	**Vincent Dandenault** *Application des réseaux bayésiens aux données multi-omiques pour l'amélioration du diagnostic de l'asthme chez les enfants d'âge préscolaire*
16h10-16h30	**Simon de Montigny** *Conceptual framework based on artificial intelligence to facilitate the integration of infectious disease modeling into public health practice*	**Marie Deprez** *Decoding Genetic Markers of Multiple Phenotypic Traits Through Biologically Constrained Genome-To-Phenome Bayesian Sparse Regression*
16h30-16h50	**Marie-Félicia Beclin** *Construction de modèles intelligents dans les données d'imagerie scanner de patients traités par benralizumab*	**Olivia Rousseau** *The Avatar method: computation of synthetic data and application in health*
16h50-17h10	**Chiara Cordier** *Reductive Discriminating Network: a new dimension reduction algorithm in context of binary classification*	**Hugo Boisaubert** *Simulation of Virtual Patient at The Operating Room*

Friday July 1rst 2022

8h30-9h00	Reception & inscription
9h00-10h00	Invited speaker 5 : **Raphaëlle Momal** (entreprise Owkin) *What deep learning in histology can bring to clinical trials design* (Amphitheater Simone Weil) Moderator: Mathieu Emilie
10h00-10h30	Coffee break
	Invited/Contributed session 3 : IA and ethics (Amphitheater Simone Weil) Moderator: Stéphane Tirard
10h30-10h50	**Jean-Michel Loubes** (Université Toulouse 3 et Artificial and Natural Intelligence Toulouse Institute) *Bias in data for Machine Learning Algorithms*
10h50-11h10	**Océane Fiant** (Université de technologie de Compiègne, laboratoire Connaissance, Organisation et Systèmes techniques) *Explainability of artificial intelligence in medicine: why contextualization matters*
11h10-11h30	**Nicolas Berkouk** (EPFL's Laboratory for Topology and Neuroscience*)* *The field of Explainable AI: creating machines to explain machines?*
11h30-11h50	**Philippe Bizouarn** *Le métier du soin virtualisé à l'heure de l'IA ?*
11h50-12h10	**Béatrice Desvergne** *L'Intelligence artificielle pour réconcilier santé personnalisée et santé publique ?*
12h10-14h00	Lunch break
14h00-17h00	Workshop « IA et R » **Tidymodels workshop** Hannah Frick **https://education.rstudio.com/trainers/people/frick+hannah/** (Amphitheater Simone Weil) Moderator: Aymeric Stamm

Abstracts

Invited speakers

10h00-11h00 **Invited speaker 1**

Grégory Nuel (LPSM-CNRS 8001, Sorbonne Université)

Bio: G. Nuel is a senior CNRS researcher of the Institute of Mathematics (INSMI) working in Laboratory of Probability, Statistics and Modeling (LPSM, CNRS 8001) at Sorbonne Université. Since 2018, G. Nuel is the head of the Stochastics and Biology Group. Throughout his career, G. Nuel has developed a genuine interest for biomedical applications in probability and statistics based on his strong theoretical background in mathematics. He is an expert in computational statistics (simulations, the expectation-maximization algorithm, Markov chain Monte Carlo techniques, etc.) and models with latent variables (Markov chains, hidden Markov models, Bayesian networks, etc.). He has a great interest for applications in bioinformatics, statistical genetics, cancer epidemiology, tropical diseases, and clinical research.

Title: «Inferring causality from a mixture of observations and interventions»

Abstract: We all know that correlation is not causality, and that cofounding factors can easily mislead us to spurious conclusions. Even when all cofounding factors are observed, we know that many causal DAGs (Directed Acyclic Graphs) structures can lead to the same likelihood; this is the so-called Markov equivalence. In simple cases, we can reduce the Markov equivalence class to a single DAG structure by using intervention experiments like clinical randomized trial. But in complex situations, even intervention experiments might not allow to infer with certainty all causal relationships. Moreover, in practice, we are often confronted to mixture of observation and intervention experiments. In this talk, we start by presenting the notion of PDAG (Partially Directed Acyclic Graphs) as a representation of the Markov equivalence class of a DAG. We will start by PDAG in the presence of observation experiments only, and then generalize to a mixture of observation and intervention experiments. We will then see how it is possible to compute the likelihood of a PDAG assuming that the underlying variables are connected to each other through generalized linear models (e.g. linear regression, logistic regression, survival, etc.). Finally, we will explain how it is possible to use the BIC criterion and MCMC (Markov Chain Monte-Carlo) in order to explore the PDAG space and derive a posterior distribution.

14h30-15h30 **Invited speaker 2**

Basile Chaix (iPLesp Sorbonne Université)

Bio: Basile Chaix is a research director at Inserm. The Nemesis team that he coordinates, created at the start of the MobiliSense project funded by the European Research Council (ERC) in 2015, examines how life environments influence health, explores the impact of transport on health (benefits and exposures associated with the different modes), and study the health effects of heat waves inside and outside the urban heat island. This work is interested in the dynamics of exposure, behavior, and health status in space and time based on fine-grained space-time referenced data. The different projects rely on a monitoring of participants with wearable sensors of location, behavior, environmental exposures, and health, and with the Eco-emo tracker smartphone application developed by the team.

Title: «The use of wearable sensors and smartphones in environment - health research»

Abstract: The objectives of the work that will be discussed are to examine how geographic life environments influence health, to explore the impact of transport on health, and to study the health effects of heat waves inside and outside the urban heat island. These studies investigate dynamics of exposures, behaviors, and health in space and time based on high frequency data. They rely on a continuous monitoring of participants with wearable sensors of location, behavior, environmental exposures, and health and real-time smartphone surveys (according to different survey strategies). Moreover, mobility surveys based on GPS data provide information on space-time budgets that are critical to interpret the data from other sensors. As an example, in the project on heat waves, participants will report their thermal discomfort, sleep quality, etc. with a smartphone; they will carry different sensors (GPS, accelerometer, heat stress measured from air and radiant temperatures, physiological sensors); and the thermal characteristics of their dwelling will be evaluated with fixed sensors. These studies allow us to investigate contextual effects on health with "momentary" analyses of sensor data. Disaggregating environmental exposures and health responses at the level of successive life segments taken as the statistical units of the analysis (trips, places visited, repeated measurements), we contextualize behaviors and health states in their immediate environment and we investigate dynamic processes that lead to unfavorable behavior and health status, with a focus on individual, environmental, and situational determinants.

9h00-10h00 **Invited speaker 3**

Rodolphe Thiébaut (INRIA - INSERM – CHU de Bordeaux - Université de Bordeaux - ISPED)

Bio: Rodolphe Thiebaut is professor of Public Health and Biostatistics at the University of Bordeaux. He is the director of the Department of Research in Public Health at the university and the department of medical information at the hospital. He is leading an Inserm/Inria research group (Statistics In Systems and Translational Medicine SISTM) devoted to statistical development applied to immunology and vaccinology. His research includes the inference and control with ODE based models, supervised and unsupervised statistical learning using longitudinal high dimensional (n<p) data. He is also leading the Graduate's program Digital Public Health that includes a Master in Public Health data science.

Title: «Analyzing transcriptomics data for understanding and predicting vaccine response in clinical trials»

Abstract: : The availability of gene expression data in vaccine trials has generated new opportunities for understanding and predicting the response to vaccine. It has led to the so-called Systems vaccinology. However, the analysis of such data is difficult because of the high dimensionality of the predictors (p) in regards of the number of available subjects (n). In this talk, I will present several methods inspired by these applications that we have developed in my team such as testing differential expression/ abundance of genes, reducing dimensions while considering geneset structures and random forest with the Frechet metrics.

14h00-15h00 **Invited speaker 4**

Stéphanie Allassonnière (Université Paris Descartes & Ecole Polytechnique)

Bio: Stéphanie Allassonnière is a professor in applied mathematics at the School of Medicine, University of Paris, PR[AI]RIE fellow and deputy director and an associate Professor in the applied Mathematics department of Ecole Polytechnique. She manages master programs and masterclasses in AI in healthcare. Her researches deal with Statistical modelling, stochastic optimization, MCMC samplers and medical data analysis in order to propose decision support systems aiming at understanding diseases response to treatments, anticipating diagnosis and therapy follow-up." She is co-founder of Sonio, a startup which provides a companion tool to support the practitioner in monitoring pregnancy, women's and children's health, and reassuring families.

Title: «Data Augmentation in High Dimensional Low Sample Size Setting Using a Geometry-Based Variational Autoencoder»

Abstract: : In this presentation, we propose a new method to perform data augmentation in a reliable way in the High Dimensional Low Sample Size (HDLSS) setting using a geometry-based variational autoencoder. Our approach combines a proper latent space modeling of the VAE seen as a Riemannian manifold with a new generation scheme which produces more meaningful samples especially in the context of small data sets. The proposed method is tested through a wide experimental study where its robustness to data sets, classifiers and training samples size is stressed. It is also validated on a medical imaging classification task on the challenging ADNI database where a small number of 3D brain MRIs are considered and augmented using the proposed VAE framework. In each case, the proposed method allows for a significant and reliable gain in the classification metrics. For instance, balanced accuracy jumps from 66.3% to 74.3% for a state-of-the-art CNN classifier trained with 50 MRIs of cognitively normal (CN) and 50 Alzheimer disease (AD) patients and from 77.7% to 86.3% when trained with 243 CN and 210 AD while improving greatly sensitivity and specificity metrics.

9h00-10h00 **Invited speaker 5**

Raphaëlle Momal (entreprise Owkin)

Bio: After a PhD in dependence networks inference and a postdoc on the study of the human gut microbiota, Raphaëlle Momal joined Owkin where she currently works on the benefits of covariate adjustment for clinical trials design, as well as gene regulatory network inference from single-cell transcriptomic data.

Title:«What deep learning in histology can bring to clinical trials design »

Abstract: Advances in deep learning allow to capture the information contained in histological images of cancer tissue. Histological images are larger than typical images processed by deep learning requiring tailored algorithms. Owkin has developed two such procedures, one relying only on information at the slide level and the other also leveraging annotations on the slide. These procedures were applied on digitized biopsies from mesothelioma (Courtiol et al. 2019) and resected HCC (Saillard et al. 2020) patients, and the resulting deep learning covariates have been validated as an independent predictor of overall survival (OS). Adjustment on prognostic covariates allows for improved precision and increased statistical power for treatment effect estimation in randomized controlled trials. In the specific setting of time-to-event outcomes, parametric and semi-synthetic simulations yield critical information on the factors impacting the reduction in sample size following covariate adjustment. We advocate for more systematic adjustment on prognostic covariates, which can lead to more efficient and more inclusive clinical trials.

14h00-17h00 **Workshop « IA et R »**

Hannah Frick (Rstudio)

Bio: Hannah Frick is a software engineer and statistician on the tidymodels team at RStudio. The tidymodels framework is a collection of packages for modeling and machine learning using tidyverse principles. She holds a PhD in statistics from the Universitaet Innsbruck and has worked in data science consultancy as well as interdisciplinary research at University College London in cooperation with Team GB Hockey.

Title: «An Introduction to tiddymodels»

Abstract: Are you a data scientist or statistician who is looking to do some machine learning? You already know why you want to split your data in training and test sets? You know which models you want to try out but don’t want to memorize the syntax details for each one? You are aware of sklearn but would prefer to work in R?

This workshop offers an introductory tour through tidymodels, a framework for modeling and machine learning using tidyverse principles. It lets you build up your workflow in clear steps with consistency, flexibility, and sensible defaults. We’ll walk through an exemplary case study to show how you can specify a range of models, bundle preprocessing and model fitting to avoid data leakage, resample your data, and tune your models to avoid overfitting

Abstracts

Invited sessions

Wednesday June 29th 2022

Invited session 1: Multi-source data integration

16h00-16h30

Anne-Sophie Jannot (University of Paris Cité and Georges Pompidou European Hospital)

Title:«Reuse of the French National Health Insurance data for patients suffering from rare diseases: the Dromos project challenges»

Abstract: Current estimates suggest that there are between 1 and 3 million rare disease patients in France. Nevertheless, data on rare diseases is still too scattered in various databases and is heterogeneous, which hinders research progress. Many avenues remain to be explored in order to better understand these diseases and thus to better diagnose them, to propose adapted and targeted treatments for each patient, to limit the medical costs, etc. The DROMOS project aims to provide a detailed description of the actual care of rare disease patients for the most frequent diagnoses on a national scale. This knowledge will help to improve the care pathways of affected people and to adapt the care offer.

Because rare diagnoses are not accurately coded in medico-admistrative database, this projet will use the National Data Bank for Rare Diseases linked to the French National Health Insurance data. This matching will make it possible to describe for the first time the care of rare disease patients on a national scale for a very large number of disease. In this presentation, I will discuss two main challenges to achieve such a goal, i.e. firstly available methods for linkage with the French National Health Insurance data and possible resulting biases and secondly methods to model clinical course from patients followed-up over different age periods and their limits.

16h30-17h00

Eric Letouzé (CRCI2NA, Inserm UMR 1307, Nantes Université)

Title: «Multi-omics data integration in cancer research»

Abstract: Tumorigenesis involves different layers of deregulation. Genomic alterations accumulate during cell divisions due to different mutational processes, and those providing a growth advantage promote the clonal expansion of tumor cells. The epigenetic and transcriptomic programs are also remodeled, allowing plasticity between more or less aggressive cell states. Here I will show how the integration of various types of omics data can unravel new oncogenic mechanisms and reveal the natural history of cancers, from tumor initiation to treatment resistance. I will also discuss how AI can help making sense of genomic data.

17h00-17h30

Erwan Drezen (entreprise CUBR)

Title:«Better data for a better science: the challenge of linking health databases»

Abstract: Any algorithm needs data. And the richer the data are, the happier the algorithm is. Especially in the field of health care, using several health data sources makes it possible to build rich patients pathways well suited for feeding algorithms. However, many issues may arise when linking heath databases, such as missing information, assessment of results or volumetry bottlenecks. We will present here a new record linkage approach and some use cases including the French National Health Insurance Information System (SNDS-

Invited session 2: Neuroscience and image analysis

10h30-11h00

Alexandre Gramfort (Inria, CEA, Université Paris-Saclay)

Title: «Machine learning without human supervision on neuroscience data»

Abstract: Understanding how the brain works in healthy and pathological conditions is considered as one of the major challenges for the 21st century. After the first electroencephalography (EEG) measurements in 1929, the 90's was the birth of modern functional brain imaging with the first functional MRI (fMRI) and full head magnetoencephalography (MEG) system. By offering noninvasively unique insights into the living brain, imaging has revolutionized in the last thirty years both clinical and cognitive neuroscience.

After pioneering breakthroughs in physics and engineering, the field of neuroscience has to face new major computational and statistical challenges. The size of the datasets produced by publicly funded populations studies (Human Connectome Project in the USA, UK Biobank or Cam-CAN in the UK etc.) keeps increasing with now hundreds of terabytes of data made available for basic and translational research.

While machine learning can offer great opportunities in this context, these datasets rarely come with strong annotations which are necessary to employ the most powerful supervised predictive models. In this talk I will present three statistical machine learning strategies applied to electrophysiological data where models are learnt without human supervision.

References:

Uncovering the structure of clinical EEG signals with self-supervised learning Banville H., Chehab O., Hyvärinen A., Engemann D., Gramfort A. (2021) Journal of Neural Engineering 18: (046020).

Combining magnetoencephalography with magnetic resonance imaging enhances learning of surrogate-biomarkers Engemann D., Kozynets O., Sabbagh D., Lemaître G., Varoquaux G., Liem F., Gramfort A. (2020) eLife 9: (e54055).

Shared Independent Component Analysis for Multi-Subject Neuroimaging Richard H., Ablin P., Thirion B., Gramfort A., Hyvärinen A. (2021) Advances in Neural Information Processing Systems 34 (NeurIPS)

11h00-11h30

Florence Forbes (Inria, Universite Grenoble Alpes, equipe Statify, Laboratoire LJK)

Title: «Simulation based inference for high dimensional inverse problems: application to magnetic resonance fingerprinting»

Abstract: A wide class of problems from medical imaging, robotics, astrophysics, economics, etc. can be formulated as inverse problems. Solving such problems generally starts by the so-called direct or forward modelling that theoretically describes how input parameters x are translated into effects y. Then from experimental observations of these effects, the goal is to find the parameter values that best explain the observed measures. Typical situations and constraints that can be encountered in practice are that 1) both direct and inverse relationships are highly non-linear; 2) the observations y are high-dimensional (eg. signals in time or spectra); 3) many such high-dimensional observations are available and the application requires a very large number of inversions; 4) the parameters x to be predicted is itself multi-dimensional with correlated dimensions. In addition, it is common that direct models are available only through mechanistic formulations that provide high-fidelity simulations of the system but only through a "black-box" poorly suited for statistical inference. The main challenge comes from the fact that the model likelihood is typically intractable and has to be estimated. These situations are referred to as likelihood-free or simulation-based inference and have received a lot of attention in recent years with momentum coming from mixing ideas from statistics and machine learning. The proposed approach is illustrated in neuroimaging and in particular with the recent concept of Magnetic Resonance Fingerprinting.

11h30-12h00

Christophe Zimmer (Institut Pasteur)

Title: «Deep learning for biomedical imaging»

Abstract: Deep learning is fueling advances and breakthroughs in a dizzying array of data-intensive scientific fields. This talk will highlight recent and ongoing work of our lab that leverages deep learning to push boundaries of biomedical imaging.
A long-standing challenge in the life sciences is to visualize biological cells at high resolution and with high throughput. Single molecule localization microscopy (SMLM) is among the most powerful and widely used super-resolution imaging methods, but is typically very slow. I will present ANNA-PALM, a computational technique based on deep learning that can reconstruct high resolution views from strongly under-sampled SMLM data, enabling considerable speed-ups without compromising spatial resolution. I will also highlight Shareloc, an online platform to facilitate the sharing and reanalysis of SMLM data, and show data on this platform can be used to increase the robustness of ANNA-PALM reconstructions. Potentially, preliminary applications to live cell super-resolution will also be shown.
Time permitting, I may also present additional projects, in which we use deep learning for medical imaging diagnostics or for characterizing antibiotic drugs.

Friday July 1rst 2022

Invited/contributed session 3: IA and ethics

10h30-10h50

Jean-Michel Loubes (PR Institut de Mathématiques de Toulouse, Université de Toulouse)

Title: «Projection to Fairness in Statistical Learning», making an estimator fair while preserving its prediction accuracy as much as possible»

Abstract: IA models have proven helpful for a large variety of medical use cases, but their instability and their lack of robustness are the Achilles' heel of modern artificial intelligence. Understanding why AI models fail is at the heart of modern research in Machine Learning. We consider the specific issues of biases in AI models who lead to bad generalization properties or some poor performance for some particular subclass of observations. We provide some definitions and ways to quantify such biases and explain some new methods to cope with such issues.

10h50-11h10

Océane Fiant (Post-doc Université de technologie de Compiègne, laboratoire Connaissance, Organisation et Systèmes techniques (Costech, UR 2223))

Title: «Explainability of artificial intelligence in medicine: why contextualization matters»

Abstract: Transparency of medical artificial intelligences is a technical, legal and ethical necessity. It has motivated the emergence of a research field on the explainability of algorithms. However, most of the proposed solutions are not oriented by a specific purpose. In fact, a majority of researches focuses on improving the comprehensibility of the models, while a minority questions the relevance of the proposed solutions for the end user. I will show in this presentation, with examples, that the development of solutions to make the results of artificial intelligences understandable to physicians should take into account the contexts these technologies are intended to integrate.

11h10-11h30

Nicolas Berkouk (post-doctorant en informatique à l'EPFL)

Title: «The field of Explainable AI: creating machines to explain machines?»

Abstract: The advent of neural networks in machine learning has brought about a paradigm shift in many fields, from applied sciences to everyday life. We are now interacting with neural networks on a daily basis: on our smartphones, while browsing the Internet or through the presence of connected objects.

Nevertheless, their internal functioning remains rather obscure, even for the scientific communities that develop them, and it is generally accepted that there is currently no satisfactory mathematical formalism to describe their learning process. Yet, would an understanding on the level of mathematics alone be sufficient? Given the huge impact of these technologies outside the laboratory, a new imperative (based on different social, legal, economic dimensions) appears: we need to produce explanations of the results of neural networks for users.

In this presentation, I will propose a preliminary analysis of the answers brought to this question by Explainable AI research, while discussing the conditions in which this very young field has been constituted.

11h30-11h50

Le métier du soin virtualisé à l'heure de l'IA ?

Philippe Bizouarn[1,2]

[1] Service d'Anesthésie-Réanimation, Hôpital Laennec, CHU, Nantes, France

[2] Laboratoire Sphere, Université de Paris, Paris, France

E-mail for correspondence: philippe.bizouarn@chu-nantes.fr

Abstract:

Michel Foucault, dans Naissance de la Clinique, pose la question suivante : « Est-il encore possible d'intégrer dans un tableau, c'est-à-dire dans une structure à la fois visible et lisible, spatiale et verbale, ce qui est perçu à la surface du corps par l'œil du clinicien, et ce qui est entendu par ce même clinicien du langage essentiel de la maladie? ». Appliqué au langage des données formalisées et analysées algorithmiquement, la question du rôle d'une IA maîtrisée par des acteurs autres que les soignants risque de conduire à un désinvestissement des tâches du soin quand le réel du travail qui se fait, auprès des patients, ne peut être reconnu par la machine car justement non formalisable. Comment, en effet, la machine algorithmique pourra rendre compte de la notion de cognition incarnée évoquée par Matthiew Crawford par laquelle les acteurs du soin agissent avec les patients. Par la réduction des corps à des données quantifiées que les soignants ne maîtriseraient plus, n'y a-t-il pas un risque d'aboutir à une forme d'objectivation de ces corps souffrants dessaisis de leur existence mondaine que le soin justement vise à valoriser ?

Keywords: Soin; cognition incarnée; objectivation; formalisation

Foucault M (1963). Naissance de la clinique. 5ème éd. (1997), Presses universitaires de France, Paris.

Crawford MB (2016). Contact. Pourquoi nous avons perdu le monde, et comment le retrouver. Ed; franç, La Découverte, Paris.

11h50-12h10

L’Intelligence artificielle pour réconcilier santé personnalisée et santé publique ?

Béatrice Desvergne

L’IA est un outil indispensable au développement de la médecine personnalisée. Grâce, entre autres, aux données génétiques propres à chacun, L’IA devrait dans un futur proche permettre de déterminer la propension de chaque individu à développer certaines maladies. Si aujourd’hui, les bénéfices de la santé personnalisée sont surtout mesurables dans les approches curatives, le volet prévention sera un élément clé de succès, notamment en termes économiques.
La santé publique, elle, a un rôle majeur de prévention à l’échelle de la population. Pourtant elle reste le parent pauvre de tous les développements qui foisonnent dans le domaine de la santé. Si la médecine personnalisée vise la prévention et fait le buzz depuis quelques années, pourquoi la santé publique, aussi responsable de la prévention, n’arrive -t-elle pas à séduire? Où sont les points de convergence entre ces deux approches ? L’IA pourrait-elle être cet outil qui va révolutionner nos pratiques de santé publique dans le domaine de la prévention ? Et pour quelles conséquences éthiques, sociétales et économiques ?

Abstracts

Contributed sessions

Abstract soumis à la conférence "IA et santé : approches interdisciplinaires"

Analysis of Cancer Dependency Maps

Antoine Bichat[1]

[1]Servier, Suresnes, France

E-mail for correspondence: antoine.bichat@servier.com

Abstract: Gene silencing is a well-known method to study how cell lines behave in the absence of one specific gene. Recently, large scale experiments managed to knock-out each one of the human genes in hundreds of cell lines (Tsherniak et al., 2017; Dwane et al., 2020). These screenings generate a large amount of data that should be analyzed with appropriate methods.
As an example, classical statistical methods on cancer dependency maps lead to the identification of TRIM8 as an essential gene in fusion-driven Ewing sarcomas. Coupled with an experimental approach, it results in the biological explanation of the dependency and a better understanding of this pediatric cancer (Seong et al., 2021).
Also, machine learning methods such as random forests or penalized linear regressions could be used to predict gene dependencies, leading to potential biomarkers for tumor vulnerabilities (Dempster, et al., 2020).
In this talk, I will present how data is generated, and how machine learning can analyze these data to perform disease understanding, target identification or even indication selection.

Keywords: Omics Data; Machine Learning; Public Data; Oncology.

Dempster, et al. Gene expression has more power for predicting in vitro cancer cell vulnerabilities than genomics. BioRxiv (2020).

Dwane, et al. Project Score database: a resource for investigating cancer cell dependencies and prioritizing therapeutic targets. Nucleic Acids Research 49.D1 (2021): D1365-D1372.

Seong, et al. TRIM8 modulates the EWS/FLI oncoprotein to promote survival in Ewing sarcoma. Cancer cell 39.9 (2021): 1262-1278.

Tsherniak, et al. Defining a cancer dependency map. Cell 170.3 (2017): 564-576.

Abstract soumis à la conférence « IA et santé : approches interdisciplinaires »

Classification semi-supervisée de séries temporelles de quaternions pour l'analyse des troubles de la marche dans la sclérose en plaques

Pierre Drouin*[1,2], Aymeric Stamm [1], Laurent Chevreuil[2], Vincent Graillot[2], Laetitia Barbin[3], Pierre-Antoine Gourraud[4,5], David-Axel Laplaud[3], Lise Bellanger[1]

[1] Laboratoire de Mathématiques Jean leray, Nantes Université, France

[2] Département recherche et développement, UmanIT, France

[3] CRTI-Inserm U1064, CIC, service de neurologie, Centre Hospitalier Universitaire, Nantes, France

[4] Nantes Université, INSERM, Centre de Recherche en Transplantation et Immunologie, UMR 1064, ATIP-Avenir, Nantes, France

[5] INSERM, CIC 1413, Pôle Hospitalo-Universitaire 11 : Santé Publique, Clinique des données, Centre Hospitalier Universitaire, Nantes, France

E-mail pour correspondance : pdrouin@umanit.fr

Résumé : L'évaluation de la marche des patients est un aspect crucial du suivi médical dans la sclérose en plaques (SEP) [1]. Dans cette étude, nous étudions la possibilité de classifier les patients en fonction de deux sources d'information : (i) la rotation de la hanche durant la marche mesurée par un capteur de mouvement sous forme de série temporelles de quaternions (QTS) et (ii) le score EDSS décrivant la gravité de la pathologie [2].

L'algorithme Quaternion Dynamic Time Warping [3] est utilisé pour mesurer la dissimilarité entre les données de marche représentées par des QTS. Il permet ainsi la généralisation aux QTS de deux méthodes semi-supervisées basée sur la classification ascendante hiérarchique : (i) une méthode par compromis nommée hclustcompro [4], (ii) une méthode par consensus nommée mergeTree [5].

Ces deux méthodes sont comparées en les appliquant aux données de marche de 27 patients atteints de SEP en utilisant leur score EDSS comme information supplémentaire. Nous évaluons leur validité sur la base de critères internes (Inertie intra groupe et indice de Dunn) ainsi que sur la pertinence clinique de leurs résultats. Les résultats de cette étude montrent la supériorité de la méthode hclustcompro dans ce contexte d'application.

Mots clés: Classicification semi-supervisées, Séries temporelles, Quaternions unitaires, Analyse de la marche, Sclérose En Plaques

Références :

[1] LaRocca NG. Impact of walking impairment in multiple sclerosis. The Patient: Patient-Centered Outcomes Research 2011;4:189–201.

[2] Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983;33:1444–1444.

[3] Jablonski B. Quaternion Dynamic Time Warping. IEEE Transactions on Signal Processing 2012;60:1174–83. https://doi.org/10.1109/TSP.2011.2177832.

[4] Bellanger L., Coulon A., Husi P. (2021) PerioClust: A Simple Hierarchical Agglomerative Clustering Approach Including Constraints. In: Chadjipadelis T., Lausen B., Markos A., Lee T.R., Montanari A., Nugent R. (eds) Data Analysis and Rationality in a Complex World. IFCS 2019. Studies in Classification, Data Analysis, and Knowledge Organization. Springer, Cham. https://doi.org/10.1007/978-3-030-60104-1_1.

[5] Hulot A, Chiquet J, Jaffrézic F, Rigaill G. Fast tree aggregation for consensus hierarchical clustering. BMC Bioinformatics 2020;21:120. https://doi.org/10.1186/s12859-020-3453-6.

Topological machine learning approach for linkage exploration in transcriptomes

Zeno Loi[1]

[1]Institut Desbrest d'Épidémiologie et de Santé Publique, Université de Montpellier, Montpellier, France
E-mail for correspondence: zeno.loi@etu.umontpellier.fr

Abstract: Modeling genes regulation networks is a major yet challenging stake to understand physiopathology. We show that genes belonging to the same regulation network have common geometrical variations when their biological function is modified by an environmental condition. This allows the first topological approach to transcriptomes analysis. First we pre-process the genes expressions data set with quantile normalization, logarithmization, removing the low expressed genes, and finally replacing the values by their z-scores. Then we use UMAP, a dimensional reduction algorithm topology preserving, to sum up the conditions observed. The local regulation networks with common behavior tend to set apart. We use a db-scan to isolate those clusters. Finally, we estimate the investment likeliness for each cluster by measuring their individual prediction performance on the studied condition. Our method provides strong inferences on which genes are implied in the cell reaction. Moreover, our method provides leads for common transcription factors among genes concerned in specific pathological situations, and thus for new therapeutic targets.

Keywords: Genes regulation network ; Machine learning; Omics data; UMAP

Abu-Jamous (2018). Clust: automatic extraction of optimal co-expressed gene clusters from gene expression data. Genome Biology.

Camara (2017). Topological methods for genomics: present and future directions. Curr Opin Syst Biol.

Dorrity (2020). Dimensionality reduction by UMAP to visualize physical and genetic interactions. Nature Communications.

Luo (2021). A topology-preserving dimensionality reduction method for single-cell RNA-seq data using graph autoencoder. Scientific Reports.

A coarse-to-fine needle extraction algorithm for the modelisation of electric field in electroporation ablation

Eloïse Inacio[1], Luc Lafitte[1], Olivier Sutter[2], Olivier Seror[2], Baudouin Denis de Senneville[1], Clair Poignard[1]

[1]Project team MONC, Univ. Bordeaux, UMR CNRS 5251, INRIA, Talence, France
[2]Interv. Radiol. Unit, Univ. Hosp. Paris Seine Saint Denis, Avicenne Hosp., APHP, Univ. Paris 13, Bobigny, France

E-mail for correspondence: eloise.inacio@inria.fr

Abstract: The aim is to provide an online estimation of the electric field applied during an electroporation ablation, following the numerical workflow in **(1)**. This minimally invasive and non-thermal ablation technique can be used on deep-seated tumors, where traditional techniques may affect vital structures. However, it requires thorough planning and evaluation, due to its inherent complexity.
To this end, we propose a novel coarse-to-fine algorithm for the extraction of needles delivering the electric field, from a single Cone Beam Computed Tomography. It is a crucial step in the computation of the electric field to evaluate the procedure **(2)**. A coarse segmentation is obtained by a modified U-Net **(3)**, trained with a patch optimization strategy and a well-suited loss function. The analytical representation is computed by a Hough transform **(4)**, completed with a voting procedure. Finally, the electric field is obtained with a standard linear static model.
The results are evaluated on 8 of 16 patients: for the coarse segmentation, we compare to the groundtruth using the Dice coefficient and for the analytical representation, the distance between the estimated and real coordinates is computed. Under two minutes on a commodity hardware, the needles are extracted with a more precise and stable algorithm than the previous coarse segmentation with thresholding.

Keywords: Deep Neural Network, Fine-object Segmentation, CBCT, Electric field distribution

(1) O. Gallinato, B. Denis de Senneville, O. Seror, C. Poignard (2019). Numerical workflow of irreversible electroporation for deep-seated tumor. Physics in Medicine and Biology.

(2) O. Gallinato, B. Denis de Senneville, O. Seror, C. Poignard (2020). Numerical Modelling Challenges For Clinical Electroporation Ablation Technique of Liver Tumors. Math. Model. Nat. Phenom.

(3) F. Isensee, et al. (2018). nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation. arXiv:1809.10486

(4) L. Shapiro, G. Stockman (2001). Computer Vision. Prentice-Hall, Inc.

Robust deconvolution of transcriptomic samples using the gene covariance structure

Bastien CHASSAGNOL[1,2,3], Pierre-Henri WUILLEMIN[1], Gregory NUEL[2], Etienne BECHT[3]

[1]LIP6 (Laboratoire d'Informatique Paris 6), Paris, FRANCE
[2]LPSM (Laboratoire de Probabilités, Statistiques et Modélisation), Paris, FRANCE
[3]Les Laboratoires Servier, Suresnes, FRANCE

E-mail for correspondence: bastien.chassagnol@upmc.fr

Abstract: Transcriptomic analyses have increasingly contributed to our understanding of the intricate biological processes involved in the emergence of auto-immune diseases or tumour-promoting environments. However, classical bulk analyses ignore the intrinsic complexity of biological samples, by averaging measurements over multiple distinct cell populations. It is therefore unclear whether a change in the gene expression between samples results from a variation of the cell type proportions or from a biological factor (**Shen-Orr and Gaujoux**, 2013).
To remove this ambiguity, deconvolution algorithms can estimate the proportions of cell populations from a bulk transcriptome using the reference transcriptome of purified cell populations. Traditionally, most approaches, including the gold standard CIBERSORT algorithm (**Abbas et al.**, 2009), retrieve the cell proportions of a mixture assuming the linear assumption that each gene expression is the sum of each cell population's contribution weighted by their corresponding relative frequency in the sample. However, none of these methods account for the transcriptomic covariance structure and address the crucial problem of co-transcriptomic expression between the genes. The first goal of our project aims at studying the impact of highly correlated structures assuming a sparse structure learnt by using the gLasso algorithm (**Friedman, Hastie, and Tibshirani**, 2008) on the performances of the canonical deconvolution algorithms using a reference-based method. Then, we will develop a new deconvolution method that integrates both the average expression and the covariance structure of the reference transcriptomic profiles to estimate cellular ratios, resolving noise effect induced by the interaction terms between gene transcripts.

Keywords: Deconvolution; Covariance; gLasso; Transcriptomic; Cell Population

Abbas, Alexander R et al. (2009). Deconvolution of Blood Microarray Data Identifies Cellular Activation Patterns in Systemic Lupus Erythematosus. PloS One, 4 (7): e6098.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani (2008). Sparse Inverse Covariance Estimation with the Graphical Lasso.. Biostatistics (Oxford, England) 9 (3): 432–41.

Shen-Orr, Shai S., and Renaud Gaujoux. (2013). Computational Deconvolution: Extracting Cell Type-Specific Information from Heterogeneous Samples. Current Opinion in Immunology 25 (5): 571–78.

Assessing Multiple Imputation of Missing Values for Robust Analysis of Telehealth Kiosk Data

Hava Chaptoukaev[1,2*], Maxime Beurey[1*], Juliette Raffort[3], Maria A. Zuluaga[1]

[1]Data Science Department, EURECOM, Sophia Antipolis, France
[2]Bodyo, Nice, France
[3]Department of Clinical Biochemistry, Nice University Hospital, Université Côte d'Azur, Nice, France
[*]Joint first authorship

E-mail for correspondence: zuluaga@eurecom.fr

Abstract: Telehealth is a promising avenue for prevention, and remote diagnosis and monitoring of diseases. However, in non-clinical contexts sensors may fail, leading to incomplete data, where features are not missing randomly, that can jeopardize subsequent analyses. In this work, we investigate if imputation schemes can improve the analysis of relationships between Blood Pressure (BP) and features collected by a stand-alone telehealth kiosk. We analyze 253 samples of 26 features, corresponding to indicators of BP, oximetry, and body composition. We use Multiple Imputation Denoising Autoencoders (MIDA) to impute missing values, with masks imitating the patterns of failing sensors in the training. In terms of mean absolute error, MIDA performs as well as mean imputation. Only MIDA is able to capture and preserve the distribution of the data. Principal Component Analysis performed on the imputed dataset suggests body weight, muscle mass and energy requirements explain 20% of data's variability, and arterial stiffness and pulse amplitude explain 15%. While Partial Least Square regression draws attention to the role of arterial stiffness, oxygen saturation, and pulse amplitude in predicting BP on the incomplete data, it also highlights the importance of body weight and muscle mass in the outcome after imputation, thus improving the analysis.

Keywords: Telehealth; Health Kiosk; Missing Values; Blood Pressure

Hotelling, H. (1993). Analysis of a complex of statistical variables into principal components. Journal of educational psychology, 24(6), 417.

Gondara L. and Wang K. (2018). Mida: Multiple imputation using denoising autoencoders. Pacific-Asia conference on knowledge discovery and data mining (pp. 260-272). Springer, Cham.

Wold S., Sjöström M., and Eriksson, L (2001). PLS-regression: a basic tool of chemometrics. Chemometrics and intelligent laboratory systems, 58(2), 109-130.

Abstract soumis à la conférence « IA et santé : approches interdisciplinaires »

Random survival forests for competing causes with multivariate longitudinal endogenous covariates

Anthony Devaux[1], Robin Genuer[1,2], Cécile Proust-Lima[1]

[1] Inserm BPH U1219, Université de Bordeaux, Bordeaux, France

[2] INRIA Bordeaux Sud-Ouest, Talence, France

E-mail for correspondence: anthony.devaux@u-bordeaux.fr

Abstract: The individual data collected throughout patient follow-up constitute crucial information for assessing the risk of a clinical event. Joint models have been proposed to compute individual dynamic predictions from repeated measures to one or two markers. However, they hardly extend to the case where the complete patient history includes much more repeated markers. We extended the random survival forest methodology to incorporate multivariate longitudinal endogenous markers. The random survival forest is composed by an ensemble of decision trees, where the subjects are recursively split into two subgroups. At each split, mixed models for the longitudinal markers are fitted and the predicted random effects are used among the others time-fixed predictors to split the subjects. The individual-specific event prediction is derived as the average over all trees of the leaf-specific cumulative incidence function. We demonstrate in a simulation study the performances of our methodology. We also applied it to predict the individual risk of dementia in the elderly according to the trajectories of cognitive functions, brain imaging markers, and general clinical evaluation. Our random survival forest extends the joint modelling methodology to predict clinical events from individual longitudinal history when the number of repeated markers is large.

Keywords: Individual prediction; Joint modeling; Random survival forests; Longitudinal modeling

Ishwaran H, Kogalur UB, Blackstone EH and Lauer MS (2008). Random survival forests. The Annals of Applied Statistics, 2(3), 841-860.

Laird NM and Ware JH (1982). Random-Effects Models for Longitudinal Data. Biometrics, 38(4), 963-974.

Abstract soumis à la conférence « IA et santé : approches interdisciplinaires »

Identification of consensus whole blood transcriptomic gene modules in patients with primary Sjögren' s Syndrome

Cheïma BOUDJENIBA[1,2,3], Etienne BIRMELE[4], Benno SCHWIKOWSKI[2], Mickaël GUEDJ[1], Etienne BECHT[1]

1 Translational Medicine, Servier, Suresnes, France

2 Laboratoire MAP5 UMR 8145, Université de Paris Cité, Paris, France

3 Computational Systems Biomedicine, Institut Pasteur, Paris, France

4 Institut de Recherche Mathématique Avancée, UMR 7501 Université de Strasbourg et CNRS, Strasbourg, France

E-mail for correspondence : etienne.becht@servier.com

Abstract:

Primary Sjögren's syndrome (pSS) is an autoimmune disease characterized by lymphoid infiltration of exocrine glands leading to dryness of the mucosal surfaces and by the production of various autoantibodies. The pathophysiology of pSS remains elusive and no treatment with demonstrated efficacy is available yet.

To better understand the system biology underlying pSS heterogeneity, we aimed at identifying Consensus Gene Modules (CGMs) summarizing the high-dimensional transcriptomic data of whole blood samples in pSS patients. We performed an unsupervised classification and identified 11 CGMs. We interpreted and annotated each of these CGMs as corresponding to cell type abundances or biological functions by using gene set enrichment analyses and transcriptomic profiles of sorted blood subsets. Correlation with independently measured cytokine levels and cell type abundances by flow cytometry confirmed these annotations.

By measuring the average expression of the CGMs on samples from clinical trials, we confirmed previously described relationships between the presence of autoantibodies, activation of the type I interferon pathway and an increased frequency of monocytes. Furthermore, we will study whether the expression of the CGMs can predict response to treatments. We believe that these CGMs will facilitate the interpretation of whole blood transcriptomic data of pSS patients.

Keywords: Precision Medicine, Sjögren's syndrome, Unsupervised learning, Integrated analysis.

Comparaisons de méthodes pour données de survie en grande dimension sur de petits échantillons : optimisation des hyperparamètres et validation.

Audrey Lavenu[1], Juliette Murris[2*], Alexis Mareau[3], Timothé Rouzé[4], Magalie Fromont[5], Valérie Gares[6] & Sandrine Katsahian[7]

[1] CIC1414 Inserm, IRMAR, Université de Rennes 1, Rennes, France

[2,3,4,7] CIC1418 Inserm, HeKA-Inria, Université de Paris, Paris, France

[5] IRMAR, Université de Rennes 2, Rennes, France

[6] IRMAR, INSA, Rennes, France

[*] Subvention de l'ANRT et Pierre Fabre (CIFRE 2020/1701)

E-mail for correspondence: audrey.lavenu@univ-rennes1.fr

Abstract: Avec l'augmentation du nombre de données sur les patients dans les domaines de l'imagerie médicale ou de la génomique, les méthodes d'analyses classiques sont souvent inadéquates dans les cas où il y a moins d'observations que de variables dans les données. Nous étudions différents critères de performance et leur estimation de la méthode Cox Boost pour analyser des données de survie en grande dimension sur petits échantillons, en termes de prédiction, de discrimination des variables pronostiques et de gain par optimisation des hyperparamètres. Nous simulons les temps de survie et de censure respectivement par des lois exponentielle et uniforme. Pour fixer le taux de censure à un taux prédéfini, nous montrons comment calculer le paramètre de la distribution de censure. En faisant varier les tailles d'effet des covariables et de l'échantillon, et le taux de variables actives, nous comparons le C de Harrell et l'importance de variable estimés par validation croisée en 2 et 5 blocs, avec trois méthodes de choix des hyperparamètres. Nous montrons la difficulté d'optimiser les hyperparamètres pour de petits échantillons, et que l'importance des variables dans le modèle utilisé ne permet pas toujours de détecter les variables simulées actives, même avec une performance correcte de prédiction.

Keywords: Méthodes d'apprentissage supervisé, Survie, Censure, Grande dimension, Simulation.

Akiba T, Sano S, Yanase T, Ohta T, Koyama M. (2019). Optuna : A next-generation hyperparameter optimization framework. Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery and data mining.

Binder H and Schumacher M. (2008). Allowing for mandatory covariates in boosting estimation of sparse high-dimensional survival models. BMC Bioinformatics. 9 :14.

Pittman J, et al. (2004). Integrated modeling of clinical and gene expression information for personalized prediction of disease outcomes. Proc Natl Acad Sci USA. 101(22) :8431-6.

Vabalas A, Gowen E, Poliakoff E, Casson A. (2019). Machine learning algorithm validation with a limited sample size. PLoS ONE. 14(11) :e0224365.

Wan F. (2016). Simulating survival data with predefined censoring rates for proportional hazards models. Stat Med. 36(5) :838-854.

Application des réseaux bayésiens aux données multi-omiques pour l'amélioration du diagnostic de l'asthme chez les enfants d'âge préscolaire

Vincent Dandenault[1,2], Simon de Montigny [2,3], Cristina Longo[2,4]

[1]Faculté de Médecine, Université de Montréal, Montréal, QC, Canada
[2]Centre de recherche du CHU Sainte-Justine, Montréal, QC, Canada
[3]École de santé publique, Université de Montréal, Montréal, QC, Canada
[4]Faculté de Pharmacie, Université de Montréal, Montréal, QC, Canada

E-mail for correspondence: `vincent.dandenault@umontreal.ca`

Abstract: Les enfants d'âge préscolaire sont particulièrement vulnérables aux symptômes d'asthme et ils sont difficiles à diagnostiquer, avec un taux presque deux à trois fois plus élevé de visites aux urgences dues aux symptômes s'apparentant à l'asthme que d'autres groupes d'âge. De plus, on en sait peu sur le bénéfice que les thérapies actuellement disponibles pourraient apporter aux enfants d'âge préscolaire souffrant d'asthme ou de respiration sifflante, ce qui souligne le besoin urgent de découvrir des biomarqueurs qui faciliteront la transition vers un traitement personnalisé. Étant donné que l'asthme est composé de divers phénotypes et endotypes moléculaires, notre approche est basée sur la recherche de nouvelles représentations naturelles de données dans l'espoir de découvrir des biomarqueurs qui pourraient être utiles dans le diagnostic et traitement de l'asthme préscolaire. En utilisant la base de données multi-omiques chez l'enfant (génome, microbiome, transcriptome et métabolome) provenant d'un projet international sur l'asthme pédiatrique, nous construisons des modèles graphiques probabilistes (PGM) qui permettent notamment d'incorporer des connaissances antérieures cliniques grâce à des distributions bayésiennes a priori. Ces approches modernes nous permettent également de visualiser graphiquement d'importantes voies causales entre les biomarqueurs omiques et les variables cliniques.

Keywords: Personalized Medicine, Machine Learning, Pediatric Asthma, OMICS, Probabilistic Graphical Models

Grad R, Morgan WJ (2012). Long-term outcomes of early-onset wheeze and asthma.. Journal of allergy and clinical immunology 130.2.

Koller D, Friedman N (2009). Probabilistic graphical models: principles and techniques. MIT press.

Licari A, et al. (2018). Asthma endotyping and biomarkers in childhood asthma.. Pediatric Allergy, Immunology, and Pulmonology 31.2.

Pijnenburg MW, Szefler S (2015). Personalized medicine in children with asthma.. Paediatric respiratory reviews 16.2.

Wainwright MJ, Jordan MI (2008). Graphical models, exponential families, and variational inference. Foundations and Trends® in Machine Learning.

Conceptual framework based on artificial intelligence to facilitate the integration of infectious disease modeling into public health practice

Simon de Montigny[1,2]

[1]CHU Sainte-Justine Research Center, Montréal, QC, Canada
[2]School of Public Health, Université de Montréal, Montréal, QC, Canada

E-mail for correspondence: simon.de.montigny@umontreal.ca

Abstract: Mathematical modeling of infectious disease transmission played a crucial role in supporting public health during COVID-19 pandemic. This crisis called for the rapid development of models and their continual maintenance and adaptation to answer a growing variety of questions as the epidemiological situation and mitigation measures evolved. Integration of real-time data in these models remains an important challenge to address with the goal of improving the relevance and impact of their previsions.
In my research program, I propose a conceptual framework that combines artificial intelligence and mathematical modeling to enable the real-time generation, calibration and simulation of infectious disease models. Using a composition modeling approach, I will design new methods and algorithms for the automated update and exploration of model structures that will help to integrate new sources of data with time-dependent granularity. Software tools implementing these innovations, to be tested in the workflow of modelers, will improve reproducibility of results and will facilitate the co-construction of mathematical models and simulation scenarios, and their prospective validation, by modelers and knowledge users. My overarching goal is to build key technologies that will enhance the contribution of mathematical modeling to public health practice, in particular for pandemic preparedness and response.

Keywords: Public health; Mathematical epidemiology; Artificial intelligence; Real-time modeling and simulation.

Becker AD, Grantz KH, Hegde ST, Bérubé S, Cummings DAT, Wesolowski A (2021). Development and dissemination of infectious disease dynamic transmission models during the COVID-19 pandemic: what can we learn from other pathogens and how can we move forward? Lancet Digital Health 3(1):e41-e50. doi: 10.1016/S2589-7500(20)30268-5.

Chretien JP, Riley S, George DB (2015). Mathematical modelling of the West Africa Ebola epidemic. eLife 4:e09186. doi: 10.7554/eLife.09186.

Muscatello DJ, Chughtai AA, Heywood A, Gardner LM, Heslop DJ, MacIntyre CR (2017). Translation of Real-Time Infectious Disease modelling into Routine Public Health Practice. Emerging Infectious Diseases 23(5):e161720. doi: 10.3201/eid2305.161720.

Gössler G, Sifakis J (2005). Composition for component-based modelling. Science of Computer Programming 55(1-3):161-183. doi: 10.1016/j.scico.2004.05.014.

Huang Y, Seck MD, Verbraeck A (2011). From data to simulation models: component-based model generation with a data-driven approach. Proceedings of the 2011 Winter Simulation Conference (WSC) 2011 Dec 11 (pp. 3719-3729). IEEE. doi: 10.1109/WSC.2011.6148065.

Schölzel C, Blesius V, Ernst G, Goesmann A, Dominik A (2021). Countering reproducibility issues in mathematical models with software engineering techniques: A case study using a one-dimensional mathematical model of the atrioventricular node. PLOS ONE 16(7):e0254749. doi: 10.1371/journal.pone.0254749.

Decoding Genetic Markers of Multiple Phenotypic Traits Through Biologically Constrained Genome-To-Phenome Bayesian Sparse Regression

Marie Deprez[1,2] , Julien Moreira[1,2], Maxime Sermesant[1,2], Marco Lorenzi[1,2]

[1] University of Côte d'Azur, Nice, France

[2] INRIA, Epione Project-Team, Valbonne, France

E-mail for correspondence: marie.deprez@inria.fr

Abstract: The applicability of multivariate approaches for joint genomic-phenomic data analysis is currently limited by the lack of scalability, and interpretability to relate findings from a biological perspective. To tackle these limitations, we present Bayesian Genome-to-Phenome Sparse Regression (G2PSR), a novel multivariate regression method based on sparse SNP-gene constraints. The statistical framework of G2PSR is based on a Bayesian neural network, were constraints on SNPs are integrated by incorporating *a priori* knowledge linking variants to their respective genes, to then reconstruct the phenotypic data in the output layer. Interpretability is promoted by inducing sparsity on the genes through variational dropout, allowing to estimate the uncertainty associated with each gene in the reconstruction task. Ultimately, G2PSR prevents multiple testing correction and assesses the combined effect of SNPs, thus increasing the statistical power in detecting genome-to-phenome associations. G2PSR effectiveness was demonstrated on synthetic and real data, with respect to state-of-the-art methods. The real data application used the Alzheimer's Disease Neuroimaging Initiative data, relating SNPs from more than 3,500 genes to clinical and multi-variate brain volumetric information. The experimental results show that our method provides accurate selection of relevant genes in dataset with large SNPs-to-samples ratio, thus overcoming limitations of current genome-to-phenome association methods.

Keywords: Bayesian, Genome, Phenome, Sparse Regression.

Blei, D. M., Kucukelbir, A., and McAuliffe, J. D. (2017). Variational Inference: A Review for Statisticians. *J. Am. Stat. Assoc.* 112, 859–877. doi:10.1080/01621459.2017.1285773

Deprez M, Moreira J, Sermesant M and Lorenzi M (2022). Decoding Genetic Markers of Multiple Phenotypic Layers Through Biologically Constrained Genome-To-Phenome Bayesian Sparse Regression. *Front. Mol. Med,* doi: 10.3389/fmmed.2022.830956

Molchanov, D., Ashukha, A., and Vetrov, D. (2017). "Variational Dropout Sparsifies Deep Neural Networks," in Proceedings of the 34th International Conference on Machine Learning - Volume 70 (JMLR.org), ICML'17, 2498–2507.

Shen, L., and Thompson, P. M. (2020). Brain Imaging Genomics: Integrated Analysis and Machine Learning. *Proc. IEEE* 108, 125–162. doi:10.1109/JPROC.2019.2947272

Construction de modèles intelligents dans les données d'imagerie scanner de patients traités par benralizumab

Marie-Félicia Béclin[1], Pierre Lafaye de Micheaux[1], Nicolas Molinari[1]

[1]IDESP, Université de Montpellier

E-mail for correspondence: marie-felicia.beclin@umontpellier.fr

Abstract: Le traitement par benralizumab est un traitement contre l'asthme dont nous voudrions prédire l'efficacité. Les questionnaires de qualité de vie ACQ (Asthma Control Questionnaire) montrent une nette amélioration pour les patients traités. L'enjeu est d'utiliser les données d'imagerie médicale afin de permettre une mesure plus objective de la réponse au traitement. La problématique est donc de prédire la réponse au traitement à l'aide des données d'imagerie scanner. Les données à disposition pour chaque patient sont les images scanner en expiration et en inspiration, les résultats des questionnaires ACQ, les données de spirométrie avant traitement, à 6 mois et à un an. Après segmentation des images, nous calculons les histogrammes. L'hypothèse est qu'un patient dont l'état s'est amélioré, aura une meilleur expiration après traitement, ce qui se traduit à l'image par des plus hautes valeurs d'HU (Hounsfield Unit), c'est-à-dire un décalage entre avant et après traitement de l'histogramme vers la droite. Nous construisons donc un modèle de régression, non pas sur des points comme usuellement mais sur les histogrammes eux-mêmes. Les histogrammes sont transformés en quantiles et nous utilisons la distance de Wasserstein. Ainsi, comme pour une régression classique, nous pouvons calculer explicitement les coefficients de notre régression.

Keywords: Données symboliques; Régression; Imagerie; Maladies respiratoires; Grande dimension

The Avatar method: computation of synthetic data and application in health

Olivia Rousseau[1], Stanislas Demuth[1,2], Chadia Ed Driouch[1,3], Nicolas Vince[1], Sophie Limou[1,4], Gilles Edan[5], Pierre-Antoine Gourraud[1,6]

[1] Nantes Université, INSERM, Centre de Recherche Translationnelle en Transplantation et Immunologie, CR2TI, Nantes, France

[2] INSERM U1119 Biopathologie de la myéline neuroprotection, Centre de recherche en biomédecine, Strasbourg, France

[3] IMT Atlantique, CNRS, LS2N, Nantes, France

[4] École Centrale de Nantes, Nantes, France

[5] Département de Neurologie, Centre Hospitalier Universitaire, Rennes, France.

[6] CHU de Nantes, INSERM, CIC 1413, Pôle Hospitalo-Universitaire 11 : Santé Publique, Clinique des données, Nantes, France

E-mail for correspondence: olivia.rousseau@univ-nantes.fr

Abstract:

For many years, clinicians and researchers have collected a huge amount of health-related data. However, more information implies a greater re-identifying risk for patients[1] and pseudonymization is not enough to overcome this data security problem. The aim of the Avatar method is to compute a synthetic dataset which conserves global characteristics of a sensitive dataset while ensuring the data security by computing re-identification metrics. Individuals are projected in a dimension reduction multidimensional space (PCA, FAMD) and then a local model is built for each sensitive individual considering his k nearest neighbors. The avatar is created randomly on this area to not easily associate a sensitive individual to its synthetic version. This method has been tested on a randomized clinical trial on multiple sclerosis treatment: the REFLEX study. A Cox analysis performed by Comi et al. gives hazard ratios 0.49 [95% CI 0.38-0.64] and 0.69 [0.54-0.87] for the two arms[2]. In comparison, the hazard ratios of a synthetic dataset computed with the Avatar method is equal to 0.44 [0.34-0.57] and 0.64 [0.50-0.82]. Synthetic databases can reproduce pseudonymous dataset analyses and can be shared to the scientific community with less re-identifying risk for patients compared to pseudonymous datasets.

Keywords: Synthetic data; Anonymization; Data protection; Multiple sclerosis

1. Rocher L and al. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. Nat. Commun.
2. Comi G and al. (2012). Comparison of two dosing frequencies of subcutaneous interferon beta-1a in patients with a first clinical demyelinating event suggestive of multiple sclerosis (REFLEX): a phase 3 randomised controlled trial. Lancet Neurol.

Reductive Discriminating Network: a new dimension reduction algorithm in context of binary classification

Chiara Cordier[1,2], Pascal Jézéquel[2], Fabien Panloup[1], Agnès Basseville[2]

[1] Laboratoire Angevin de Recherche en Mathématiques, Faculté des Sciences, Angers, France

[2] Unité Omiques et Data Science, Institut de Cancérologie de l'Ouest, Angers/Nantes, France

E-mail for correspondence: chiara.cordier@ico.unicancer.fr

Abstract:

In oncology, machine learning (ML) is implemented for personalized medicine to predict patient response to treatment and select optimal treatment accordingly. ML models are built from high-dimensional omics data, and therefore require a dimension reduction step to avoid "n<<p" ML issue. For now, dimension reduction algorithms are either unsupervised or not always dedicated to keep important information for following-step classification[1].

In that purpose, we developed the Reductive Discriminating Network (ReDiN). This new deep learning algorithm reduces dimension with a focus on binary classification. As in Generative Adversarial Networks[2], two neural networks learn simultaneously: one reduces data, the other assigns a score to each reduced sample to evaluate its class. These networks are optimized so that the Wasserstein distance between distributions of the two reduced data classes is maximized.

Random forests (RF) were used afterward to evaluate ReDiN on final prediction performance. Several synthetic datasets mimicking biological data were used to identify dataset key characteristics leading to performance variations. With uncorrelated-variable datasets, ReDiN improved RF prediction from 30% initial error to 10%. Increasing correlated variable proportion in datasets led to error classification augmentation from 10% to 45%. Accordingly, we used network inference and graph neural networks to further improve ReDiN performances.

Keywords: dimension reduction; deep learning; binary classification; Wasserstein distance

1. Review of dimension reduction methods. S. Nanga, A.T. Bawah, B.A. Acquaye, M.-I. Billa, F.D. Baeta, N.A. Odai, S.K. Obeng, A.D. Nsiah. Journal of data analysis and information processing, 2021

2. Generative adversarial nets. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio. Advances in neural information processing systems, 2014

SVP-OR: Simulation of Virtual Patient at The Operating Room

Hugo Boisaubert[1] , Christine Sinoquet[1] , Corinne Lejus-Bourdeau[2,3]

[1] Laboratoire des Sciences du Numérique de Nantes (LS2N / UMR CNRS 6004), Nantes Université, France

[2] Laboratoire Expérimental de Simulation de Médecine Intensive (LE SiMU), Nantes Université, France

[3] Service d'anesthésie-réanimation, CHU de Nantes, Nantes Université, France

E-mail for correspondence: hugo.boisaubert@univ-nantes.fr

Abstract:

Since 2012, the High Authority for Health has imposed a new paradigm: "Never a first time on the patient" which has led to the development of medical simulation. We present SVP-OR, a generator of reactive scenarios designed to provide digitally assisted training for interns in anaesthesiology (SVP-OR users). To simulate a virtual patient (VP), we predict the short-term evolution of the VP's multivariate time series grown so far, each time an action is triggered by the intern or the virtual medical team. The prediction problem is tackled as a case-based reasoning approach: the evolution of the VP is computed from the real patients showing some region of their histories similar to the VP's recent history. A grammar built from real patients' action traces drives the consistent scheduling of the SVP-OR user-induced actions and of the medical team's virtual actions. Our contributions are the design of a contextualized multidimensional pattern recognition approach, and a vast comparative study of dissimilarity measures focused on short time series. We evaluated four variants of our generic SVP-OR framework. We showed in all cases that SVP-OR is able to generate on-the-fly realistic predictions.

Keywords: Computer-assisted Medical Training; Virtual Patient; Operating Room ; Multivariate Time Series Prediction; Event Trace;

Abstracts

Posters

Modélisation d'anesthésie

par représentation synthétique de traces opératoires

Hugo Boisaubert[1] , Christine Sinoquet[1] , Corinne Lejus-Bourdeau[2,3]

[1] Laboratoire des Sciences du Numérique de Nantes (LS2N / UMR CNRS 6004), Nantes Université, France

[2] Laboratoire Expérimental de Simulation de Médecine Intensive (LE SiMU), Nantes Université, France

[3] Service d'anesthésie-réanimation, CHU de Nantes, Nantes Université, France

E-mail for correspondence: hugo.boisaubert@univ-nantes.fr

Abstract:

Plus d'un million d'opérations chirurgicales avec anesthésie générale sont réalisées chaque jour dans le monde. Les CHU ont obligation d'enregistrer les profils anesthésiques de patients. Chaque profil anesthésique comporte la trace des actions médicales et d'anesthésie appliquées au patient, et la série temporelle décrivant l'évolution de ses paramètres physiologiques. A partir des traces observées pour un type d'opération chirurgicale, nous souhaitons construire une représentation synthétique capable de capturer la variabilité des opérations, à un niveau d'abstraction contrôlé.

Nous construisons d'abord le graphe orienté exact où tout chemin représente une réalisation de l'opération chirurgicale d'intérêt. Une représentation synthétique est ensuite obtenue par agglomération d'arcs, selon un processus ascendant hiérarchique. Nous contrôlons le niveau d'abstraction en informant le processus d'agglomération sur la similarité des régions de séries temporelles associées aux arcs. Une agglomération d'arcs entraîne aussi une agglomération de séries temporelles en une série consensus. Un alignement hiérarchique ascendant de type Dynamic Time Warping permet de calculer ce consensus. Une version stochastique approchée est proposée.

Les applications sont nombreuses: simulation de patient virtuel au bloc opératoire, anonymisation de données temporelles complexes, interprétation de la variabilité des opérations chirurgicales.

Keywords: Anesthésie, Représentation synthétique, Abstraction, Anonymisation, Analyse de la variabilité

Abstract soumis à la conférence « IA et santé : approches interdisciplinaires »

Diverse data in multiple sclerosis improves machine learning performance for short-term evolution of disease trajectories, lessons from EPIC cohort.

Antoine Lizée[1, 2], Adam Santaniello[1], Stanislas Demuth[3, 4], Bruce A C Cree[1], Jorge R Oksenberg[1], Stephen L Hauser[1], Sergio Baranzini[1], Riley Bove[1], Pierre-Antoine Gourraud[1, 3]

[1] Department of Neurology, UCSF Weill Institute for Neurosciences, San Francisco, CA, United States.

[2] Data Community, Alan, Paris, France

[3] INSERM U1064, Centre de Recherche Translationnelle en Transplantation et Immunologie, CR2TI, Université de Nantes, Nantes, France

[4] INSERM U1119 Biopathologie de la myéline et neuroprotection, Université de Strasbourg, Strasbourg, France

E-mail for correspondence: stanislas.demuth@chru-strasbourg.fr

Abstract:

Introduction: Multiple sclerosis (MS) is a typical multifaceted neurological disease requiring the integration of several modalities of data (clinical, imaging, biological, genetics) for diagnosis and management. Machine learning could significantly support tackling the characteristic unpredictability of MS.

Methods: We analyzed a prospective longitudinal cohort of 589 patients with MS from the EPIC study at the University of California San Francisco (UCSF), totaling 3,456 yearly clinical visits. We applied a supervised classification machine learning framework, considering each yearly clinical visit as a datapoint. Features were clinical, imaging, genetic and quality of life (QoL) data at the index visit. The label was the short-term clinical evolution at the next visit.

Results: The best predictive performance was achieved by the SVM classifier with all longitudinal features included in the model (AUROC: 0.782 ±0.003), which outperformed 2 expert neurologists (AUROC: 0.54 and 0.57). The addition of QoL and timed functional metrics to the usual feature set of MS research significantly improved predictive performances (AUROC increase of 0.04 and 0.05, respectively).

Conclusion: Machine learning offers significant short-term predictive performances in MS follow-up. The impact of variable selection on performance shows that input data are more important that the type of algorithms used in the prediction.

Keywords: Precision medicine, Artificial intelligence, prediction, machine learning, multiple sclerosis.

Abstract soumis à la conférence « IA et santé : approches interdisciplinaires »

Proposal of a universal digital format for reference disease trajectories in chronic medical conditions: the segment of reference.

Stanislas Demuth[1, 2], Chadia Ed-Driouch[1, 3], Olivia Rousseau[1], Gilles Edan[4], Cédric Dumas[3], Jérôme De Sèze[2], Pierre-Antoine Gourraud[1], for the PRIMUS consortium

[1] INSERM U1064, Centre de Recherche Translationnelle en Transplantation et Immunologie, CR2TI, Université de Nantes, Nantes, France

[2] INSERM U1119 Biopathologie de la myéline et neuroprotection, Université de Strasbourg, Strasbourg, France

[3] CNRS UMR 6004 Nantes Information Sciences Laboratory, Université de Nantes, France

[4] Department of neurology, Rennes University Hospital, Rennes, France

E-mail for correspondence: stanislas.demuth@chru-strasbourg.fr

Abstract:

Introduction: Unlike model-driven approaches of precision medicine, contextualization is a data-driven approach yielding prognostic insight by describing the disease trajectories of matching patients in reference datasets. Building reference datasets for a given condition requires a universal format to conciliate multiple data sources.

Methods: We considered 3 multiple sclerosis studies datasets, processed by the PRIMUS consortium, with different designs: 2 randomized clinical trials (the time-to-event REFLEX study and the fixed follow-up ADVANCE study) and a real-world dataset of the Strasbourg OFSEP center.

Results: Patient follow-ups were converted into a tree model in json format. Iterative algorithms extracted segments of reference (SORs) between all possible timepoints triplets within a homogenous therapeutic sequence: (1) past history for a given time window, (2) matching variables at the index timepoint and (3) endpoint variables at a given prediction horizon. Our contextualization algorithm calls all matching SORs given a set of variables chosen by the user.

Discussion: Compared to machine learning, the sets of matching and endpoint variables and their stratifications may be customized in the contextualization request made by the clinical user depending his reasoning.

Conclusion: We propose a universal SOR format to support data-driven approaches of precision medicine where massive data is available.

Keywords: Big data; Multiple sclerosis; Digital health; Contextualization, Decision support.

Metastatic breast cancer PET image registration for longitudinal lesion segmentation

Constance Fourcade[1,2]
Ludovic Ferrer[4,5] *PhD* Noémie Moreau[2] Gianmarco Santini[2] *PhD*
Aislinn Brennan[2] Caroline Rousseau[3,5] *MD PhD* Marie Lacombe[5] *MD*
Vincent Fleury[5] *MD* Mathilde Colombié [5] *MD* Pascal Jézéquel[4,5] PhD
Mario Campone[3,5] *MD, PhD* Mathieu Rubeaux[2] *PhD* Diana Mateus[1] *PhD*

[1]Ecole Centrale de Nantes, LS2N, UMR CNRS 6004, Nantes, France
[2]Keosys Medical Imaging, Saint Herblain, France
[3]University of Nantes, CRCINA, INSERM UMR1307, CNRS-ERL6075, Nantes, France
[4]University of Angers, CRCINA, INSERM UMR1307, CNRS-ERL6075, Angers, France
[5]ICO Cancer Center, Nantes - Angers, France
E-mail for correspondence: `constance.fourcade@ec-nantes.fr`

Abstract:

Metastatic breast cancer presents a poor prognosis and requires constant monitoring. Hence, images are regularly acquired to assess the evolution of tumors over time. When interpreting images, physicians follow guidelines such as Position Emission tomography Response Criteria In Solid Tumors (PERCIST). However, these guidelines tend to focus only on some lesions representing tumor burden, since assessing global tumor evolution is challenging and time-consuming.

Using our longitudinal metastatic breast cancer Position Emission Tomography (PET) images acquired in the context of the study EPICURE$_{\text{seinmeta}}$, we aim to propagate lesion segmentation manually performed by experts on baseline acquisitions to the follow-up images.

On 110 baseline/follow-up pairs of images, we expanded the patient-specific registration method MIR-RBA [Fourcade et al., 2022] to improve the segmentation of lesions on longitudinal data. Thus, we integrated segmentation information in the loss function and as additional network channels of the registration pipeline.

We showed that the segmentation loss alone does not refine segmentations, while integrating lesion information in the registration network significantly improved the segmentation accuracy in term of both Dice score and detection rate.

As a conclusion, we improved longitudinal lesion segmentation in the context of metastatic breast cancer incorporating segmentation information at different levels of an image registration pipeline.

Clinical trial: NCT03958136

Research sponsor: FEDER-FSE n°PL0015129 (EPICURE)

Keywords: PET; Image registration; Segmentation propagation; Breast cancer

Fourcade C., Ferrer L., Moreau N., Santini G., Brennan A., Rousseau C., Lacombe M., Fleury V., Colombié M., Jézéquel P., Campone M., Rubeaux M., Mateus D.(2022). Deformable Image Registration with Deep Network Priors: a Study on Longitudinal PET Images. arXiv preprint arXiv:2111.11873.

On the use of CNNs for Covid'19 Detection using Lung CT Images

Sonia Hamoutene[1] , Assia Kourgli[2]

[1] Faculté de Génie Electrique, USTHB, Alger, Algérie

E-mail for correspondence: assiakourgli@gmail.com, akourgli@usthb.dz

Abstract: In this paper, a deep-learning-based approach, namely convolutional neural networks (CNNs) are used in order to classify COVID-19 and normal (healthy) computer tomographic (CT) scans images. The image database was extracted from (Chowdhury 2020 and Rahman 2021). It contains 300 images in total, including 200 for the training base, 80 for the validation and 20 for the tests. First, we designed our own CNNs architecture. The results obtained were unconvincing, but this was to be expected given the number of images on which our model was trained. Indeed, this number is insufficient for deep learning. This is why we hypothesized that working with transfer learning algorithms would improve our classification by using pre-trained models. Thus, we used a pre-trained model offered by the Keras library, the VGG16 that offers the possibility to be used directly or modified according to the desired application. . The latter has demonstrated its capabilities by greatly improving our classification. Thus, as was the case during the previous classification, performance measures such as accuracy, precision, recall, F1-score and AUC of the ROC curve were taken into account to assess this classification. Our next objective is to quantify the degree of lung damage in percentage.

Keywords: Deep learning, CNNs, CT scan images, Covid'19 detection

Chowdhury M., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M., Mahbub, Z. and al. (2020). Can AI help in screening Viral and COVID-19 pneumonia? (IEEE, Éd.) IEEE Access, **8**, pp. 132665 - 132676.

Kogilavani S. V., Prabhu., Sandhiya , Sandeep Kumar , Subramaniam , Karthick A. , Muhibbullah M., Banu S. Imam S. (2020). COVID-19 Detection Based on Lung Ct Scan Using Deep Learning Techniques", Computational and Mathematical Methods in Medicine, vol. 2022, Article ID 7672196, 13 pages, 2022. https://doi.org/10.1155/2022/7672196

Rahman , T., Khandakar, A., Yazan , Q., Tahir , A., Kiranyaz, S., Kashem, **S., Chowdhury, M.** (2021). Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images **.** Computers in Biology And Medecin, **132,** 104319.

Silva, P., Luz E., Silva G., Moreira G., Silva R., Lucio D., Menotti D. (2020). COVID-19 detection in CT images with deep learning: A voting-based scheme and cross-datasets analysis, Informatics in Medicine Unlocked, **20**, 100427,ISSN 23529148,https://doi.org/10.1016/j.imu.2020.100427.

Smeulders, A. W. (2000). Content-based image retrieval at the end of the early. (IEEE, Éd.) IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1349-1380.

Analyse en Composantes Principales de séries temporelles de quaternions unitaires: application à l'analyse de la marche chez les patients atteints de sclérose en plaques

Klervi LE GALL[1] Lise BELLANGER[1] David-Axel LAUPLAUD[2] Aymeric STAMM[1]

[1]Laboratoire de Mathématique Jean Leray, UMR CNRS 6629, Nantes Université, France
[2]CR2TI, INSERM U1064, CHU de Nantes, Nantes Université, France

E-mail for correspondence: `Klervi.Legall@univ-nantes.fr`

Abstract:
L'analyse de la marche est devenue un levier important dans la compréhension et la prise en charge de certaines pathologies comme la sclérose en plaques. Les dispositifs permettant cette analyse sont de plus en plus nombreux et variés tels que les capteurs de mouvement. Dans le cadre d'une collaboration de recherche, l'équipe de recherche ALEA du Laboratoire de Mathématiques Jean Leray avec l'entreprise nantaise UmanIT développent le dispositif eGait (brevet en cours de dépôt ; eGait, 2021). Ce dispositif permet de construire un biomarqueur appelé signature de marche (SdM). La SdM caractérise la rotation de la hanche d'un individu au cours d'un cycle de marche moyen. Les rotations sont représentées par des quaternions unitaires. La SdM fournit une mesure quantitative de la démarche à un moment donné. Les objectifs de ce travail sont (i) d'adapter des méthodes exploratoires de type Analyse en Composantes Principales (ACP) aux SdM, (ii) de construire des SdM synthétiques à l'aide d'une approche mêlant ACP et distances entre SdM observées. Nous illustrerons ce travail à l'aide d'un échantillon de 30 patients atteints de sclérose en plaques issu d'une étude menée en collaboration avec l'équipe de neurologie du CHU de Nantes.

Keywords: Analyse en Composantes Principales; Séries temporelles de Quaternions unitaire; Marche; Sclérose en plaques; Données Synthétiques

Abstract soumis à la conférence « IA et santé : approches interdisciplinaires »

A Hidden Markov Model for the surveillance of infections by the bovine viral diarrhoea virus in cattle

Aurélien Madouasse[1], Annika van Roon[2], Gerdien van Schaik[2, 3], Jörn Gethmann[4], Jude Eze[5], Maria Guelbenzu-Gonzalo[6], Christine Fourichon[1]

[1] INRAE, Oniris, BIOEPAR, 44300, Nantes, France

[2] Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, Utrecht, Netherlands

[3] Royal GD, Deventer, The Netherlands

[4]Institute of Epidemiology, Friedrich-Loeffler-Institute, Südufer 10, 17493 Greifswald, Germany

[5]Scotland's Rural College, Kings Buildings, West Mains Road, Edinburgh, EH9 3JG, United Kingdom

[6]Animal Health Ireland, Carrick on Shannon, Co. Leitrim, Ireland

E-mail for correspondence: aurelien.madouasse@oniris-nantes.fr

Abstract: In order to control major cattle infectious diseases, control programmes that rely on longitudinal test data for surveillance are commonly implemented. Infection by the bovine viral diarrhoea virus (BVDV) is widespread and has important consequences on animal welfare and farm profitability. BVDV surveillance programmes implemented in different areas use different tests and testing intervals. This heterogeneity creates difficulty for trade because herds certified as free from infection in different programmes may have different probabilities of hosting infected animals. Our aims were to develop a model that estimates herd-level probabilities of infection from heterogeneous data and to compare the predicted probabilities of infection in 5 countries: France, Germany, Ireland, the Netherlands and Scotland. We describe a Bayesian Hidden Markov Model (Madouasse et al. 2022) that uses historical data to predict a probability of infection for each herd in the surveillance programme. It was not possible to compare predicted probabilities of infection between France where an ELISA antibody test was used and the other countries that used PCR or antigen tests, because it was not possible to derive priors for test characteristics relating to a common latent status. In the other 4 countries, the model predicted high posterior probabilities of infection freedom.

Keywords: Veterinary; Surveillance; Bayesian inference; Stan

Madouasse, A., Mercat, M., van Roon, A., Graham, D., Guelbenzu, M., Santman Berends, I., van Schaik, G., Nielen, M., Frössling, J., Ågren, E., Humphry, R., Eze, J., Gunn, G., Henry, M.K., Gethmann, J., More, S.J., Toft, N., Fourichon, C. (2022). A modelling framework for the prediction of the herd-level probability of infection from longitudinal data. Peer Community J. 2, e4. https://doi.org/10.24072/PCJOURNAL.80

Abstract soumis à la conférence « IA et santé : approches interdisciplinaires »

Integrating large-scale genetic data revealed a non-HLA immune locus associated with kidney graft survival

Vincent Mauduit[1], Axelle DURAND[1], Rokhaya BA[1], Martin MORIN[1], Clémence PETIT[1,2], Pierre-Antoine GOURRAUD[1], Nicolas VINCE[1], Sophie LIMOU[1,3]

[1] Université de Nantes, CHU Nantes, Inserm, Centre de Recherche en Transplantation et Immunologie, UMR 1064, ITUN, Nantes, France

[2] CHU de Nantes, Service de Néphrologie et d'immunologie clinique, Nantes, France

[3] Ecole Centrale de Nantes, Département de Mathématiques, Informatique et Biologie, Nantes, France

E-mail for correspondence: vincent.mauduit@etu.univ-nantes.fr

Abstract

Introduction: Donor-recipient mismatches in HLA genes have been associated with a poorer kidney graft survival. However, HLA mismatches alone do not explain long-term graft function decline. We ran a genome-wide association study (GWAS) on a large monocentric cohort of kidney transplanted recipients in order to characterize genetic factors associated with kidney graft loss beyond HLA.

Methods: The KiT-GENIE DNAbank comprises 1778 recipients of European ancestry for kidney transplants performed in Nantes since 2000. First, we ran a Cox proportional hazards model for time-to-failure (defined as return to dialysis or preemptive retransplantation) and time-to-rejection events. Using the residuals of the Cox model, we then led the genetic association study on >8.2M polymorphisms (SNPs) with the FaST-LMM program.

Results: No statistically significant association was identified for biopsy-proven rejection. However, we identified a significant association for graft failure near the IFNAR1 gene (p=2.53x10-8, OR 1.59) which is involved in immune response signaling pathways.

Conclusion: Our preliminary analyses in a large homogeneous monocentric cohort revealed a non-HLA immune genetic factor associated with graft failure. Further validation in non-European recipients and external cohorts will be performed. In addition, we plan on investigating the impact of donor genetic factors and donor-recipient mismatches.

Keywords: Transplantation, Genetics, Association Study, Survival Analysis, HLA

Sellarés et al. (2012). Understanding the causes of kidney transplant failure: the dominant role of antibody-mediated rejection and nonadherence

Sasaki et al. (2010). The HLA-matching effect in different cohorts of kidney transplant recipients: 10 years later.

Lippert et al. (2011). FaST linear mixed models for genome-wide association studies

Poster 9 Jenny Moreno

Numériser, échantillonner et analyser la modalisation des émotions de l'aidant du malade de cancer au travers de sa parole

Jenny MORENO
Laboratoires PREFICS et LLING CNRS

Résumé

La figure de l'aidant du malade de cancer peut faire apparaître des rôles et des liens sociaux et affectifs divers selon le contexte où l'aidé y vit. Les dernières statistiques de l'Institut National du Cancer ont estimé que cette maladie représente la première cause de mortalité chez l'homme, et la deuxième chez la femme, en France métropolitaine. En 2018 ont été recensés 204 600 cas chez l'homme et 177 400 chez la femme. Une telle situation nous interpelle et nous conduit à observer quels sont les dispositifs (structures) publics d'accompagnement et quelles sont les configurations qui se construisent autour de l'aidance des malades des cancers en France.

L'intérêt scientifique de ce travail est, en se situant au carrefour des SHS et de la Santé, de faire une étude exploratoire sur les émotions— au regard des postulats de l'analyse du discours — à partir du recueil de corpus d'entretiens menés auprès des aidants naturels de malades des cancers. Le traitement des données recueillies sera fait à l'aide du logiciel de lexicométrie Iramuteq. L'interprétation de ces corpus permettra de dévoiler les représentations linguistiques et socioculturelles que l'aidant a de *soi,* de l'aide qu'il accomplit et plus largement, de l'aidance du malade.

Mots clés : aidance, maladie de cancer, lexicométrie, analyse du discours, émotions.

Poster 10 Jean-Michel Nguyen

DEEPNIC : une nouvelle technologie pour transformer chaque variable en une image pour l'apprentissage profond.

Jean-Michel Nguyen, CRCINA, Nantes

Mathieu Brunner, Ecole Centrale, Nantes

Su Ruan, LITIS, Rouen

Transformer chaque variable en une image permettrait d'appliquer les méthodes du DeepLearning utilisant des CNNs sur les données tabulaires. Les méthodes publiées transforment chaque base de données en repositionnant chaque variable de cette dans une image. La position des variables est définie par ses coefficients de corrélation avec les autres variables. Nous avons développé une méthode permettant de transformer chaque variable en une image, la ROP-Image. Cette technologie utilise une nouvelle famille d'informations statistiques, les NICs (Nguyen Information Criteria, *Bioinformatics 2021*), développées à partir d'un nouveau type de forêts aléatoires (Random Forest of Perfect Trees), constitués exclusivement d'arbres de classification parfaites issus du modèle ROP (Regression Optimized). Pour chaque ROP-Image, l'intensité de chaque pixel est définie par la probabilité d'obtenir une classification parfaite en présence de la variable. La position de chaque pixel est définie par le nombre d'inputs et leurs pondérations utilisés dans les neurones artificiels du modèle ROP. 54675 images sont produites selon cette technologie, à partir des valeurs de l'expression de gènes issues d'une puce Affymetrix® (résistance des cancers du sein au paclitaxel, GSE22513) et 30 autres images à partir de la base de données Wisconsin (*WorldCist 2022*). Les analyses des ROP-Images par les CNNs ont commencé.

Test non-paramétrique pour la comparaison d'échantillons de données de séquençage à haut débit en cellule unique.

A. Ozier-lafontaine[1,2], F. Picard[3] and B. Michel[1,2]

[1]Ecole Centrale de Nantes
[2]Laboratoire de Mathématiques Jean Leray
[3]LBMC, ENS-Lyon

Abstract: La technologie du séquençage à haut débit en cellules uniques permet désormais de quantifier l'expression des gènes d'un échantillon à l'échelle de la cellule unique, encodés sous la forme de matrices de comptages contenant des milliers d'observations (les cellules) et des dizaines de milliers de variables (l'expression des gènes).

L'analyse de ces données nécessite le développement de méthodes adaptées à leur complexité ainsi qu'à leur taille. Une problématique courante consiste à comparer la distribution de l'expression d'un ou plusieurs gènes entre deux ou plusieurs conditions (ex: Contrôle, Traitement 1, Traitement 2). Nous proposons d'aborder cette question en nous inspirant des travaux de [**?**] [**?**] et [**?**].

Nous développons un test non-paramétrique qui s'appuie sur une méthode de classification supervisée. Dans ce cadre, la statistique est solution d'un problème d'optimisation inspiré de l'analyse discriminante de Fisher à noyaux (KFDA), dont la résolution permet d'identifier les axes discriminants les conditions testées. Cette méthode a donc un autre avantage : celui de fournir une méthode de représentation des observations combinée au test.

Notre procédure fait partie de la famille des tests à noyaux, dont le plus connu est le test par Maximum Mean Discrepancy (MMD), qui s'appuie sur une distance entre l'espérance des représentants des distributions dans un espace de Hilbert autoreproduisant (RKHS). L'approche KFDA se distingue par la prise en compte de la structure de covariance des observations dans ce RKHS, ce qui s'avère central dans le domaine de la génomique en cellules uniques pour mieux prendre en compte la variabilité biologique.

Bien qu'étudié théoriquement (sa distribution asymptotique sous l'hypothèse nulle est établie), le coût computationnel élevé de la méthode a considérablement limité son utilisation, ce coût étant lié à la diagonalisation de l'opérateur de covariance des représentants. Dans ce travail nous développons une version efficace du test par KFDA grâce à une méthode de Nystrom tout en garantissant le niveau du test. Nous illustrons notre méthode en l'appliquant à la comparaison de la distribution d'expression des gènes dans des populations de cellules.

A visualization model for health-related big data: personalized patient contextualization after solid organ transplantation

Olivia Rousseau[1], Estelle Geffard[1], Axelle Durand[1], Nicolas Vince[1], Sophie Limou[1,2], Pierre-Antoine Gourraud[1,3]

[1] Nantes Université, INSERM, Centre de Recherche Translationnelle en Transplantation et Immunologie, CR2TI, Nantes, France

[2] École Centrale de Nantes, Nantes, France

[3] CHU de Nantes, INSERM, CIC 1413, Pôle Hospitalo-Universitaire 11 : Santé Publique, Clinique des données, Nantes, France

E-mail for correspondence: olivia.rousseau@univ-nantes.fr

Abstract:

Several models have previously been developed to predict/monitor kidney transplanted patients outcomes[1,2]. Here, we propose to leverage the large amount of collected clinical data to offer an individualized model in a precision medicine perspective. With this in mind, we collect data on 800 individuals in the KTD-innov research project, which are our patients of interest (POI); and use more than 7000 individuals from the DIVAT cohort as our population of reference (POR). Based on a factorial analysis for mixed data[3], we use 24 variables to contextualize POIs in the POR as a proof of concept. We then select three methods to subdivide the POR to better represent the POI: (1) filtering method, clinicians can choose values or scale of values for each variable; (2) nearest neighbors' approach, clinicians select the POR size and (3) clustering approach, with a hierarchical clustering on principal component method[4], clusters are pre-defined and chosen according to the POI. Finally, biological values of POI, such as creatine levels, can be displayed through quantile regression or generalized additive models performed on POR values; comparable to growth curves in child health record. The next step is to apply this method with omic data such as genomics or transcriptomics.

Keywords: Big data; Contextualization; Kidney transplantation; Precision Medicine

1. Foucher Y and al. (2010). A clinical scoring system highly predictive of long-term kidney graft survival. Kidney International.
2. Chi D Chu and al. (2021). The Kidney Failure Risk Equation for Prediction of Allograft Loss in Kidney Transplant Recipients. Kidney Medecine.
3. Pagès J. (2004). Analyse factorielle de données mixtes. Revue de Statistiques Appliquées.
4. Josse J. (2010). Principal Component Methods - Hierarchical Clustering - Partitional Clustering: Why Would We Need to Choose for Visualizing Data?

Abstract soumis à la conférence "IA et santé : approches interdisciplinaires"

On automatizing the computation of reproduction numbers for deterministic compartmental models of infectious diseases

Alexandre Simard[1,2], Vincent Dandenault[3], Bruno Curzi-Laliberté[4], Michalis Famelis[2], Marios-Eleftherios Fokaefs[4], Simon de Montigny[3,5]

[1]Mila, Québec Artificial Intelligence Institute, Montréal, QC, Canada
[2]Department of Computer Science and Operations Research, Université de Montréal, Montréal, QC, Canada
[3]CHU Sainte-Justine Research Center, Montréal, QC, Canada
[4]Department of Computer Engineering and Software Engineering, Polytechnique Montréal, Montréal, QC, Canada
[5]School of Public Health, Université de Montréal, Montréal, QC, Canada

E-mail for correspondence: `alexandre.simard.5@umontreal.ca`

Abstract: The basic reproduction number $\mathcal{R}_0$ and the effective reproduction number $\mathcal{R}_t$ are important characteristics of infectious disease models. In the case of deterministic compartmental models, comprising groups of susceptible (S), infected (I) and recovered (R) persons, these numbers can be computed analytically by deriving formulas (for simple models) or numerically by using the next-generation matrix method (for more complex models). They can also be approximated using certain simulation outputs. Overall, setting up and validating the procedure to compute reproduction numbers is a time-consuming task that depends on the model's structure.
We propose a straightforward approach to compute $\mathcal{R}_0$ and $\mathcal{R}_t$ by generating and simulating a new version of the model with three generations for I and R compartments that track separately a primary case, the secondary cases that it produces and the subsequent cases. This procedure be can readily generalized to a large variety of model structures and it is very simple to validate since the transformed model contains an explicit representation of reproduction numbers.
This new approach, based on standard operations on the structure of models, will be applied in a meta-modeling framework which will help automatize the creation, calibration and validation of infectious disease models.

Keywords: Reproduction numbers; Deterministic models; Mathematical epidemiology; Generations; Model structure.

Diekmann O., Heesterbeek J.A.P. and Roberts M.G. (2010). The construction of next-generation matrices for compartmental epidemic models. Journal of the Royal Society, Interface, 7, 873-885. https://doi.org/10.1098/rsif.2009.0386.

Heesterbeek, J.A. (2002). A Brief History of R0 and a Recipe for its Calculation. Acta Biotheor 50, 189-204. https://doi.org/10.1023/A:1016599411804.

Thron C., Mbazumutima V., Tamayo L.V. et al. (2021). Cost effective reproduction number based strategies for reducing deaths from COVID-19. Journal of Mathematics in Industry 11, 11. https://doi.org/10.1186/s13362-021-00107-6.

van den Driessche P. and Watmough J. (2008). Further notes on the basic reproduction number. Brauer F., van den Driessche P. and We J. (eds) Mathematical Epidemiology. Lecture Notes in Mathematics, vol 1945. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-78911-6_6.

var den Driessche P. (2017). Reproduction numbers of infectious disease models. Infectious Disease Modelling, 2(3), 288-303. https://doi.org/10.1016/j.idm.2017.06.002.

Consult your mail to vote for the best poster

on balotilo.org

before Thursday June 30th at 01:30 pm

Allasonnière	Stéphanie	Université Paris Cité
Babin	Etienne	Oniris-INRAE
Basseville	Agnès	ICO
Beclin	Marie-Félicia	IDESP
Bellanger	Lise	LMJL, Nantes Université
Berkouk	Nicolas	EPFL
Bézier	Clément	BIO LOGBOOK
Bichat	Antoine	Servier
Bihouéé	Audrey	Nantes Université
Bizouarn	Philippe	CHU Nantes
Boisaubert	Hugo	Nantes Université
Bolut	Clémence	CNRS-Laboratoire Restore-IRIT
Boudjeniba	Cheïma	Servier, Institut Pasteur
Brehon	David	Air Pays de la Loire
Brisseau	Nadine	Oniris
Carlier	Thomas	CHU Nantes
Causeur	David	Institut Agro
Chabeau	Lucas	UMR 1246 SPHERE / Sêmeia
Chaix	Basile	Inserm
Chaptoukaev	Hava	EURECOM
Charpentier	Eric	Institut du thorax
Chassagnol	Bastien	Les Laboratoires Servier
Codet	Marie	Bio Logbook
Cordier	Chiara	LAREMA et ICO
Coulon	Arthur	Université de Tours
Courtois	Emeline	Inserm
Dabo	Sophie	Université de Lille
Dandenault	Vincent	Université de Montréal
De Almeida Braga	Cédric	Nantes Université
De Montigny	Simon	Université de Montréal

De Visme	Sophie	INSERM
Dehman	Alia	Aviwell
Demuth	Stanislas	Université de Strasbourg
Deprez	Marie	INRIA
Desvergne	Béatrice	Université de Lausanne
Devaux	Anthony	Bordeaux Population Health
Drézen	Erwan	CUBR
Drouin	Pierre	UmanIT / LMJL
Ducrot	Lucas	Sorbonne Université
Dussap	Bastien	Université Paris-Saclay
Emily	Mathieu	Institut Agro Rennes Angers
Falco	Antonio	Universidad Cardenal Herrera
Forbes	Florence	Inria
Frick	Hannah	RStudio
Gaignard	Alban	CNRS
Galassi	Francesca	Univ Rennes 1
Galboni	Adama	AMU
Galharret	Jean-Michel	LMJL
Gares	Valérie	INSA Rennes
Gaultier	Aurélie	Nantes Université
Gouasmia	Abdelkrim	Université de Tebessa
Gourraud	Pierre-Antoine	Nantes Université
Gramfort	Alexandre	Inria, Université Paris-Saclay
Grenouilloux	Armelle	CFV
Guedj	Mickael	Nanobiotix
Guyomarch	Béatrice	CHU de Nantes
Hêche	Félicien	HEIG-VD
Hussein	Burhan Rashid	Inria Rennes
Idjahnine	Hanane	Centre hospitalo-universitaire
Inacio	Eloïse	INRIA/Université de Bordeaux

Jaber	Afaf	Univeristé de Picardie
Jannot	Anne-Sophie	Université Paris Cité / APHP
Kourgli	Assia	USTHB
Laporte	Fabien	Institut du Thorax INSERM
Lavenu	Audrey	Université de Rennes 1
Le Corff	Sylvain	Institut Polytechnique de Paris
Le Gall	Klervi	LMJL
Le Meur	Nolwenn	EHESP
Lebrun	Ewen	Santeclair
Lefebvre	Alexandra	Sorbonne Université
Lehebel	Anne	INRAE
Letouzé	Eric	CRCI2NA
Loi	Zeno	CHU Montpellier
Madouasse	Aurélien	UMR BIOEPAR, INRAE - Oniris
Manet	Ghislain	CD76
Martinroche	Guillaume	Université de Bordeaux / Inria
Mateus	Diana	Centrale Nantes
Meurée	Cédric	INRIA
Michel	Bertrand	Ecole Centrale de Nantes
Molinari	Nicolas	Université de Montpellier
Momal	Raphaëlle	Owkin
Montalibet	Virginie	Université de Bordeaux-IMB
Montalvo Zulueta	Nigreisy	Université Paris Cité
Moreno	Jenny	Nantes Université
Nuel	Gregory	CNRS / Sorbonne Université
Ouazzani	Kevin	Elsan
Ozier-Lafontaine	Anthony	Nantes Université
Perez	Daniel	École normale supérieure
Perthame	Emeline	Institut Pasteur
Proia	Frédéric	Université d'Angers

Proust-Lima	Cécile	INSERM
Rolland	Jakez	LS2N/Bio Logbook
Rousseau	Olivia	CR2TI, Nantes Université
Saint Pierre	Philippe	Institut de Mathématiques de Toulouse
Saulnier	Tiphaine	Université de Bordeaux
Savy	Nicolas	Institut de Mathématiques de Toulouse
Ségalas	Corentin	Université de Paris / Inserm CRESS
Seznec	Bruno	Paris 7
Sinoquet	Christine	Nantes Université
Stamm	Aymeric	CNRS
Tea	Illa	Nantes Université
Thiebaut	Rodolphe	Université de Bordeaux
Tirard	Stéphane	Nantes Université
Vigneau	Evelyne	Oniris
Yazzourh	Sophia	Institut de Mathématiques de Toulouse
Zimmer	Christophe	Institut Pasteur

Coming to MSH

FROM DOWNTOWN:

At the station «Commerce», take the bus line 54 in the direction of «Saupin», or the bus line C3 in the direction of «Bd de Doulon». Get off at the station «Tbilissi». Walk to the big traffic circle, take the second exit on the right «Allée Jacques Berque».

The cost of one-hour-valid ticket is 1,70€. There are some vending machines at the stop «Commerce». You can also buy a book of 10 tickets which costs 15,60€. Breakdown ticket available from the driver on board the bus for 2 €.

FROM THE TRAIN STATION:

Upon the arrival at the main station, take the South exit (Sortie GareSud) and walk in the direction of the canal of Saint-Félix. Go down the «quai Malakoff» following the descending order of the numbers until the big traffic circle. Take the second exit on the right «Allée Jacques Berque».
It will take you about ten minutes to walk.

FROM THE AIRPORT:

Get to the city center by the airport shuttle bus (navette Aéroport) in 20 minutes. The final stop of the shuttle is «Commerce». There is one bus every 30 minutes. From there, take the bus line as indicated in the paragraph «FROM DOWNTOWN».

◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇◇

See the public transport map on page 70
For detailed bus schedules and maps please visit TAN *http://www.tan.fr*

Website airport *https://www.nantes.aeroport.fr/fr*

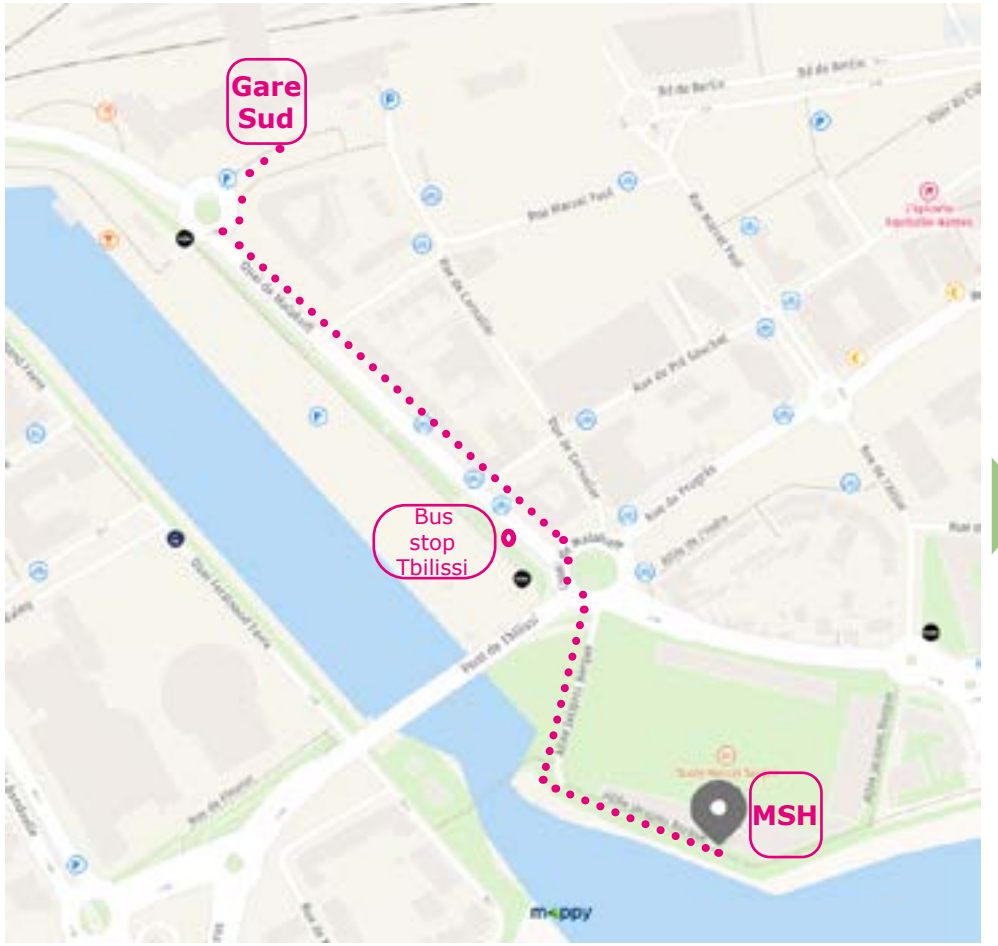
Website MSH https://msh-ange-guepin.univ-nantes.fr/

Maison des Sciences de l'Homme
Ange Guépin
5, allée Jacques Berque
44021 Nantes cedex 1

Map 1: Public transport map

For detailed bus schedules and maps please visit TAN http://www.tan.fr

Map 2: Maison des Sciences de l'Homme (MSH)



Walking distance


https://www.nantes-tourisme.com/fr
https://www.lestablesdenantes.fr
Crédits : Johanna Buguet

Crédits : Ryan Klaus